



Entrar

Cadastrar

Esta é uma cópia de uma conversa entre ChatGPT e Andréino.
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 the columns and their data types. Then, I'll apply the necessary filtering conditions to create the cleaned dataset (`df_filtered`). Let's begin.

The dataset contains 30,130 entries and 34 columns, with key fields including:

- `admission_period`: The period when a student was admitted.
- `enrollment_type`: The type of enrollment (e.g., "Normal", "Exempted").
- `situation`: The final academic status of the student (e.g., "Approved", "In Progress").

Now, I'll filter the dataset based on the given conditions:

1. Keep only records where `admission_period` is between 2006.1 and 2019.2.
2. Remove entries where `enrollment_type` is "Exempted".
3. Exclude records where `situation` is "In Progress".

I'll proceed with these filters and store the cleaned data in `df_filtered`.

The cleaned dataset (`df_filtered`) now contains 10,065 records while maintaining all 34 columns. It excludes:

- Admission periods outside 2006.1 to 2019.2.
- Students with "Exempted" enrollment type.
- Students still "In Progress".

Now, you can provide a specific analytical question, and I'll extract insights strictly based on this filtered dataset.

Using appropriate statistical models (e.g., ANOVA, multiple regression), evaluate whether the type of admission affects the number of semesters required for graduation while controlling for confounding factors such as socioeconomic status and program choice.

To evaluate whether the type of admission affects the number of semesters required for graduation, while controlling for confounders like socioeconomic status and program choice, we will:

1. Define Key Variables

- Independent Variable (IV): `admission_form` (Type of admission).
- Dependent Variable (DV): `dropout_period_count` (Number of semesters until graduation).
- Confounders:
 - `quota_form` (Proxy for socioeconomic status, as quotas often apply to disadvantaged groups).
 - `curriculum_code` (Represents different academic programs).

2. Filter for Graduated Students

- Since we are analyzing the number of semesters until graduation, we will filter only those with `exit_form = "Graduate"`.

3. Apply a Statistical Model

- ANOVA (Analysis of Variance): To check if the type of admission significantly affects the number of semesters.
- Multiple Regression: To control for confounding factors (`quota_form`, `curriculum_code`).

Let's perform this analysis now.

Key Findings from the Analysis

1. ANOVA Results (Effect of Admission Type on Graduation Time)

- F-statistic: 100.36 (very high), indicating significant variation.
- p-value: 1.26×10^{-121} , which is far below 0.05, confirming that admission type significantly impacts the number of semesters required for graduation.

2. Multiple Regression Results (Controlling for Confounders)

- $R^2 = 0.128$, meaning the model explains 12.8% of the variance in graduation time.
- Key Effects of Admission Type:
 - SISU students graduate ~2.61 semesters earlier than the baseline (Entrance Exam).
 - Reentry students graduate ~8.67 semesters earlier (likely due to previous coursework).
 - Entrance Exam / ENEM students graduate ~0.29 semesters later than the baseline.
 - Transfers & Reoptions show no significant delay.
 - Judicial Decision admissions do not significantly affect graduation time.
- Quota Effects:
 - School Race quotas reduce graduation time by 1.81 semesters (significant at $p < 0.001$).
 - Other quota types show minimal or no significant effect.
- Program Choice (Curriculum Code):
 - Students from curriculum 2017 take ~2.24 more semesters to graduate, suggesting curricular differences.

Conclusion

- Yes, the type of admission significantly affects graduation time.
- SISU and Reentry students graduate faster, while traditional Entrance Exam students take longer.
- Quotas have mixed effects, with School Race quotas leading to slightly shorter graduation times.
- Program choice strongly influences graduation time.

Let me know if you need further analysis or visualizations! 📊 []

Envie uma mensagem para o ChatGPT

