

Checkpoint #1 — 2025_2

K-Means em IMDB Top 250 (2 Notebooks)

Integrantes do grupo:

- Thiago Almança da Silva — RM 558108
- _____ — RM _____
- _____ — RM _____
- _____ — RM _____

Repositório GitHub (contendo os 2 notebooks .ipynb):

<https://github.com/<usuario>/<repositorio>>

Arquivos esperados no repositório:

- notebooks/01_scrape_and_kmeans_synopsis.ipynb
- notebooks/02_kmeans_all_features.ipynb
- data/imdb_top250_raw.csv
- data/imdb_top250_k5_synopsis.csv
- data/imdb_top250_k5_allfeatures.csv
- README.md

Observações:

- Execute primeiro o Notebook 1 para gerar a base raspada.
- Em seguida, rode o Notebook 2 e compare os modelos (M1 vs M2).
- Escreva os insights/conclusões finais e justifique o melhor modelo.

Prazo de envio (Teams): 14/09/2025 às 23:59 — America/Sao_Paulo