

## EXERCÍCIOS DE APOIO

Apenas para praticar. Não vale nota.

Faça o download do arquivo disponível - [COM350\\_Univesp\\_ExerciciosApoioSemana4.ipynb](https://ava.univesp.br/bbcswebdav/xid-4969630_1) ([https://ava.univesp.br/bbcswebdav/xid-4969630\\_1](https://ava.univesp.br/bbcswebdav/xid-4969630_1)), faça a importação dele no seu Google Colab para facilitar o processo.

Exercícios:

- 1 - Importe as bibliotecas Pandas, Numpy, Seaborn, Matplotlib e Sweetviz.
- 2 - Importe os dados de um arquivo CSV com informações de filmes da Marvel, que está disponível em: <https://www.kaggle.com/rachit239/mcu-complete-dataset> (<https://www.kaggle.com/rachit239/mcu-complete-dataset>).
- 3 - Usando a biblioteca Sweetviz, verifique qual foi o valor de orçamento (budget) que aconteceu mais vezes nos filmes da Marvel, e indique qual a porcentagem de vezes que esse valor foi determinado para os filmes.
- 4 - Usando uma correlação de Pearson, verifique no database da Marvel, quais são as variáveis que têm maior correlação positiva.
- 5 - Entre no Portal Brasileiro de Dados Abertos e baixe o arquivo CSV da Relação de Empreendimentos de Geração (<https://dados.gov.br/dataset/relacao-de-empreendimentos-de-geracao> (<https://dados.gov.br/dataset/relacao-de-empreendimentos-de-geracao>)). Informe qual é a licença atribuída a esse dado.
- 6 – Importe somente as primeiras 500 linhas desse arquivo CSV que foi baixado. Observe que como esse arquivo é de origem brasileira, precisaremos utilizar o parâmetro encoding='iso-8859-1' na função read\_csv do Pandas.
- 7 - Com os dados de Empreendimentos de Geração, indique qual é o atributo que mais tem valores nulos e elimine-o.
- 8 – Com os dados de Empreendimentos de Geração, indique qual é o atributo que tem valores nulos na menor quantidade e elimine as linhas que têm esses atributos com valor nulo.
- 9 - Com os dados de Empreendimentos de Geração, crie uma nova coluna chamada “mdaPotenciaOutorgadaMW”, baseada na coluna “mdaPotenciaOutorgadaKW”. Observe que a nova coluna é em Megawatt e a anterior Kilowatt.
- 10 - Com os dados de Empreendimentos de Geração, na coluna datOperação, deixe apenas a data nos valores, eliminando a informação sobre horas.
- 11 - Com os dados de Empreendimentos de Geração, crie uma nova coluna a partir da coluna “sigTipoGeracao”, em que os valores categóricos (qualitativos) serão substituídos por valores numéricos.

MOSTRAR GABARITO