**ATIVIDADE DA AULA 11/09/2025**

import pandas as pd

import unidecode

data = r'dados\_nao\_tratados.csv'

df = pd.read\_csv(data)

df.columns = (

    df.columns.str.strip()

              .str.upper()

              .str.replace(' ', '\_')

              .str.replace('[^0-9A-Z\_]', '', regex=True)

)

# IMPORTEI O DOCUMENTO CSV E PADRONIZEI AS COLUNAS

df['NOME'] = (

    df['NOME']

    .str.replace(r"\(.\*?\)", "", regex=True)

    .str.strip()

    .str.upper()

    .apply(unidecode.unidecode)

)

# PADRONIZEI OS DADOS DA COLUNA NOME

df['ATIVO'] = df['ATIVO'].str.upper().map({

    'SIM': True, 'TRUE': True, 'YES': True,

    'NÃO': False, 'NAO': False, 'FALSE': False, 'E': False

})

# TROQUEI OS DADOS PARA NUMERO PARA TIRAR STRINGS

df['ID'] = pd.to\_numeric(df['ID'], errors='coerce')

df['IDADE'] = pd.to\_numeric(df['IDADE'], errors='coerce')

df['DATA\_INSCRICAO'] = pd.to\_datetime(df['DATA\_INSCRICAO'], errors='coerce', dayfirst=True)

df['NOTA'] = pd.to\_numeric(df['NOTA'], errors='coerce')

df['DATA\_INSCRICAO'] = df['DATA\_INSCRICAO'].fillna('DESCONHECIDO')

df['IDADE'] = df['IDADE'].fillna('DESCONHECIDO')

df['NOTA'] = df['NOTA'].fillna('DESCONHECIDO')

# TIRA OS DADOS NULOS COLOCANDO DESCONHECIDO NELES

df.to\_csv('dados\_tratados.csv', index=False)