

BIBLIOTECA PANDAS




LEITURA


 ***pandas.read_csv("local do arquivo", encoding = "UTF-8", sep = ";")*** retorna um dataframe

 ***pandas.read_excel("local do arquivo.xlsx", ...)*** transforma o arquivo em um dataframe

 ***após o sep, header = linha*** dataframe inicia da linha especificada

 ***após o sep, usecols = ["coluna1", "coluna2"]*** retorna os dados das colunas especificadas










 ***após o sep, nrows = número*** retorna a quantidade de linhas especificada

 ***após o sep, sheet_name = número ou "nome"*** abre a aba especificada pelo número


 ***variável = pandas.ExcelFile("local do arquivo")*** retorna dataframe mais específico de excel

 ***variável.parse("nome da aba")*** retorna dataframe somente da aba especificada

COMANDOS

-  ***dataframe.head()*** retorna as colunas e as 5 primeiras linhas
-  ***dataframe.shape*** retorna a quantidade de linhas e colunas
-  ***variável.sheet_names*** retorna o nome das abas do arquivo
-  ***variável["coluna"]*** retorna um objeto com os dados da coluna
-  ***variável.coluna*** retorna um objeto com os dados da coluna
-  ***variável[["coluna1", "coluna2"]]*** retorna um dataframe com os objetos da coluna
-  ***variável.loc[linha inicial: linha final] ou [[1, 3, 7]]*** seleciona as linhas desejadas
-  ***variável.loc[linha : linha, "coluna"]*** seleciona as linhas desejadas da coluna especificada
-  ***variável.loc[linha : linha, ["coluna1", "coluna2"]] ou [linha : linha, "coluna" : "coluna"]***
seleciona as linhas desejadas da coluna especificada

 ***variável.iloc[0 : 6, 2 : 5]*** o iloc funciona com os números para linhas e colunas mas o final é -1

 ***variável.iloc[[2, 6], 2 : 5]*** intervalos não contínuos