

LISTA DE EXERCÍCIOS

CSV, Manipulação de dados

Antônio Pedrozo

- 1 • **Escreva um script em python, para ler um arquivo CSV de nome “Roupas_venda”, e printar o nome de suas colunas.**

```
import pandas as pd
```

```
arquivo=pd.read_csv('roupas_venda.csv')  
print(arquivo.columns)
```

2 • Encontre os 3 erros no código abaixo e escreva para que serve esse programa.

```
import pandas as pd

arquivo=pd.read_excel(o certo é _csv)("MICRODADOS_ENEM_ESCOLA.csv", encoding="ANSI",sep=";")
RS=arquivo.loc[arquivo["SG_UF_ESCOLA"]=="RS"]
RS_2013=RS.loc[RS["NU_ANO"]==2013 (fechar cochetes)
(faltou botar os números das linhas junto ao nome (arquivo['Name'] [0:10]) eu acho

media=RS_2013["NU_MEDIA_CN"].mean

print(media)
```

O programa serve para fazer a pontuação média no enem de 2013 no RS.

3 • Complete o código abaixo no espaço amarelo. (Pode usar o arquivo pokemon_data.csv que está na nossa pasta de aula)

```
import PAndas as pd

# Importa arquivo
pokemon=pd.read_csv("pokemon_data.csv")

# lista com maior HP
poke_HP=pokemon.sort_values(["HP"],ascending=False).reset_index(drop=True)

# dois maiores HP
poke1=poke_HP.loc[0]
poke2=poke_HP.loc[1]
```

4 • Neste contexto, por que usar reset_index?

```
import [REDACTED] as pd

# Importa arquivo
pokemon=pd.read_csv("pokemon_data.csv")

# lista com maior HP
poke_HP=pokemon.sort_values(["HP"], [REDACTED]=False).reset_index(drop=True)

# dois maiores HP
poke1=poke_HP.loc[[REDACTED]]
poke2=poke_HP.loc[[REDACTED]]
```

A função para redefinir o índice do objeto de série fornecido. Saída: como podemos ver na saída, a Series. reset_index() função redefiniu o índice do objeto Séries fornecido para o padrão.

Neste código ela é utilizada para criar uma lista com índices que vão ser utilizados nas últimas linhas do programa/código.

5 • Escreva a diferença entre a função **loc** e **iloc**, dando um exemplo de cada.

```
import pandas as pd

# Importa arquivo
pokemon=pd.read_csv("pokemon_data.csv")

#Associa pokemons
poke1=pokemon.loc[pokemon["Name"]=="Ivysaur"]
poke2=pokemon.iloc[[2]]
```

Illoc é versátil, já que podemos selecionar colunas pela nomenclatura delas
Já o iloc depende dos números dos índices, o que torna mais rara sua utilização.

Ex: A função loc é necessário que informe o nome do item.
Na iloc é preciso o Índice onde ele se localiza.