

Documentação Missão Prática

Autor: Italo Gabriel de Almeida Costa

Instituição: Estácio

Descrição do Projeto

Aplicação que demonstra a manipulação e limpeza de dados utilizando o **pandas** em Python.

1. Ler o conteúdo do CSV:

```
colunas = ['ID', 'Duration', 'Date', 'Pulse', 'Maxpulse', 'Calories']
dados = pd.read_csv('database.csv', sep=';', header=None, names=colunas,
encoding='utf-8')
```

2. Verificar os dados importados:

```
print(dados)
```

3. Informações gerais sobre o conjunto de dados:

```
print(dados.info())
```

4. Primeiras 10 linhas do arquivo:

```
print(dados.head(10))
```

5. Últimas 10 linhas do arquivo:

```
print(dados.tail(10))
```

6. Criar uma cópia do conjunto de dados:

```
dados_copia = dados.copy()
```

7. Substituir valores nulos da coluna 'Calories' por 0:

```
dados_copia['Calories'].fillna(0, inplace=True)
```

8. Substituir valores nulos da coluna 'Date' por '1900/01/01':

```
dados_copia['Date'].fillna('1900/01/01', inplace=True)
```

9. Transformar a coluna 'Date' em datetime:

```
dados_copia['Date'].replace('1900/01/01', pd.NA, inplace=True) dados_copia['Date'] =
```

```
pd.to_datetime(dados_copia['Date'], format='%Y/%m/%d', errors='coerce')
```

10. Verificar a transformação da coluna 'Date':

```
print(dados_copia)
```

11. Corrigir o valor "20201226":

```
dados_copia['Date'].replace('20201226', '2020/12/26', inplace=True)
```

12. Reaplicar a transformação da coluna 'Date':

```
dados_copia['Date'] = pd.to_datetime(dados_copia['Date'], format='%Y/%m/%d',  
errors='coerce')
```

13. Imprimir o conjunto de dados após as correções:

```
print(dados_copia)
```

14. Remover registros com valores nulos:

```
dados_copia.dropna(inplace=True)
```

15. Imprimir o dataframe final:

```
print(dados_copia)
```