



Nome: Wallace Felipe Tavares Moreira

Matrícula: 202109237331

Universidade: UNIVERSIDADE ESTÁCIO DE SÁ

Curso: Desenvolvimento Full Stack

Campus: Jardim América – Itaguaí/RJ

Disciplina: Nível 3 – RPG0033 - TRATANDO A IMENSIDÃO DOS DADOS

Semestre Letivo: Quinto Semestre

Missão Prática | Tratando a Imensidão dos Dados

Limpeza e preparação de dados para análise usando Python e Pandas

Contextualização

Como Analista de Dados, você recebeu um novo projeto com a responsabilidade de preparar um conjunto de dados para análise. Este documento descreve o processo passo a passo utilizado para limpar e preparar os dados, tornando-os aptos para mineração e análise de dados.

Prática

- **Procedimentos**

1. Leitura e Verificação Inicial

- Leitura do conteúdo do CSV fornecido.
- Atribuição dos dados lidos a uma variável.
- Verificação das informações gerais sobre o conjunto de dados.
- Impressão das primeiras e últimas N linhas do arquivo.

2. Criação de uma cópia dos dados

- Criação de uma nova variável contendo uma cópia dos dados originais.

3. Tratamento da coluna 'Calories'

- Substituição de valores nulos da coluna 'Calories' por 0.

4. Tratamento da coluna 'Date'

- Substituição de valores nulos da coluna 'Date' por '1900/01/01'.

5. Transformação da coluna 'Date' para datetime

- Utilização do método 'to_datetime' para transformar os dados da coluna 'Date' em datetime.

6. Correção de formatos específicos na coluna 'Date'

- Correção de valores específicos como "20201226" na coluna 'Date' para o formato datetime.

7. Remoção de registros com valores nulos na coluna 'Date'

- Remoção de registros onde a coluna 'Date' possui valores nulos.