



Estácio

UNIVERSIDADE ESTÁCIO DE SÁ
RIO DE JANEIRO - POLO SANTA CRUZ
2024

DOCUMENTAÇÃO DE PROJETO

DESENVOLVIMENTO FULL STACK

MUNDO 5 – NÍVEL 3

RPG0033 – TRATANDO A IMENSIDÃO DOS DADOS

Aluna: Tamires de Souza Alves

Matrícula: 202203974271

Objetivos da Prática

Este projeto tem como objetivos principais:

1. **Ler um arquivo CSV** usando a biblioteca Pandas (Python).
2. **Criar um subconjunto de dados** a partir de um conjunto existente.
3. **Configurar o número máximo de linhas** a serem exibidas na visualização de um conjunto de dados.
4. **Exibir as primeiras e últimas “N” linhas** de um conjunto de dados.

Introdução

A atividade “Tratando a Imensidão dos Dados” tem como objetivo proporcionar aos alunos a experiência prática de manipulação e limpeza de dados utilizando a biblioteca Pandas em Python.

Este projeto foi desenvolvido para simular o trabalho de um Analista de Dados, onde os participantes aprenderão a ler, filtrar, e transformar conjuntos de dados reais.

Contextualização

Os alunos receberão um conjunto de dados no formato CSV, que pode ser obtido de fontes abertas na web. O foco da atividade é preparar esses dados para análises posteriores, garantindo que estejam organizados e livres de inconsistências. Este processo inclui a substituição de valores nulos, transformação de tipos de dados e remoção de entradas problemáticas.

Metodologia

A atividade é dividida em microatividades com o objetivo de orientar os alunos em cada etapa do processo de manipulação de dados:

1. **Leitura do CSV:** Importação do arquivo e exibição dos seus conteúdos
2. **Subconjuntos de Dados:** Criação de variáveis para armazenar subconjuntos com colunas selecionadas.
3. **Configuração de Exibição:** Ajuste das configurações de visualização na biblioteca Pandas.

4. **Visualização:** Impressão das primeiras e últimas linhas do DataFrame.
5. **Análise e Limpeza de Dados:** Identificação e substituição de valores nulos, transformação de tipos de dados e remoção de registros inválidos.

Resultados

Ao final da atividade, espera-se que os alunos sejam capazes de:

- Manipular e analisar dados utilizando a biblioteca Pandas de forma eficaz.
- Compreender a importância do tratamento de dados para a análise de informações.
- Aplicar as técnicas aprendidas em projetos futuros de análise de dados.

Conclusão

A atividade oferece uma abordagem prática e estruturada para o aprendizado de manipulação de dados, preparando os alunos para desafios reais na análise de dados. Com esta experiência, os que a concluíram estarão mais bem que equipados para trabalhar em projetos que exigem habilidades analíticas e técnicas.