

SENAI A. JACOB LAFER

ANA B. CASA GRANDI, GABRIEL G. SILVA, GIULIA S. VASCONCELOS,
LARA A. RIBEIRO, MATHEUS F. NASCIMENTO.

PANDAS

Santo André

2024

ANA B. CASA GRANDI, GABRIEL G. SILVA, GIULIA S. VASCONCELOS,
LARA A. RIBEIRO, MATHEUS F. NASCIMENTO.

PANDAS

Trabalho apresentado ao curso de
análise e desenvolvimento de sistemas

Orientadores: Valmir e Ignácio.

Santo André

2024

Sumário

Introdução	4
O que é?	4
Como funciona?	4
Estruturas de Dados.....	4
Como é utilizada?.....	4
Instalação	5
Conclusão	5
Referencias	5

Introdução

Com o crescimento exponencial da quantidade de dados disponíveis, a capacidade de manipular e analisar esses dados de forma eficiente tornou-se crucial. O Pandas é uma biblioteca essencial para cientistas de dados e analistas, oferecendo uma vasta gama de ferramentas para lidar com dados de maneira eficiente e eficaz.

O que é?

Pandas é uma biblioteca de software para a linguagem de programação Python, destinada à manipulação e análise de dados. Foi criada por Wes McKinney em 2008 e tem se tornado uma das bibliotecas mais populares no campo da ciência de dados.

Como funciona?

Pandas fornece estruturas de dados como Series (1-dimensional) e DataFrame (2-dimensional), que facilitam a manipulação de dados de várias fontes, como arquivos CSV, bancos de dados SQL e planilhas Excel. Com Pandas, você pode realizar operações como leitura, filtragem, agregação e visualização de dados com poucas linhas de código.

Estruturas de Dados

- **Series:** Uma Series é uma estrutura de dados unidimensional que pode conter dados de qualquer tipo (inteiros, strings, floats, etc.).
- **DataFrame:** Um DataFrame é uma estrutura de dados bidimensional, semelhante a uma tabela de banco de dados ou uma planilha do Excel.

Como é utilizada?

Pandas é amplamente utilizada em várias etapas do processo de ciência de dados:

- **Análise de Dados Exploratória (EDA):** Permite a inspeção visual e estatística dos dados para encontrar padrões, detectar anomalias e testar hipóteses.
- **Limpeza e Preparação de Dados:** Inclui a manipulação de dados faltantes, transformação de dados, remoção de duplicatas e integração de múltiplas fontes de dados.
- **Manipulação de Séries Temporais:** Oferece ferramentas poderosas para trabalhar com dados temporais, como reamostragem, mudança de frequência e cálculo de diferenças.

Instalação

Instalar o Pandas é um processo simples que pode ser realizado de várias maneiras, dependendo do seu ambiente de desenvolvimento e das suas necessidades específicas. Porém, o método mais utilizado é através do pip. Sendo ele o gerenciador de pacotes padrão para Python e a maneira mais comum de instalar bibliotecas Python. Para instalar o Pandas usando pip, abra o terminal (ou prompt de comando) e execute o seguinte comando: **pip install pandas**.

Conclusão

Pandas é uma ferramenta essencial para qualquer pessoa que trabalhe com dados em Python. Sua capacidade de manipular e analisar dados de maneira eficiente a torna indispensável para cientistas de dados, analistas de dados e desenvolvedores. Com uma vasta gama de funcionalidades e uma comunidade ativa, Pandas continua a evoluir e se adaptar às necessidades da ciência de dados moderna.

Referencias

- Documentação Oficial do Pandas
- McKinney, W. (2012). **Python for Data Analysis**. O'Reilly Media.
- Alura