

Aluno: Rafael Pereira Cândido

Matrícula: 2221134

Curso: Ciência da Computação

Proposta de Solução

Objetivo do Projeto:

- Elaborar relatório analítico interativo utilizando um pipeline de Big Data, PySpark e Streamlit combinados.

Qualificações do Projeto:

- Dados em grande volumetria/contexto real.
- Inclusão de metadados de modelagem na pipeline.

Descrição do Projeto:

Motivação:

O setor de combustíveis no Brasil desempenha um papel fundamental na economia e na vida cotidiana dos cidadãos. A variação dos preços dos combustíveis pode impactar diretamente o orçamento das famílias, a competitividade das empresas e a formulação de políticas públicas. Diante disso, é essencial realizar uma análise abrangente dos dados disponíveis para compreender os padrões e tendências dos preços dos combustíveis ao longo do tempo e em diferentes regiões do país.

Objetivo:

O objetivo deste projeto é realizar uma análise exploratória dos dados históricos de preços dos combustíveis no Brasil, identificando padrões, tendências e insights relevantes. Além disso, pretende-se desenvolver um modelo preditivo para estimar os preços futuros dos combustíveis com base nos dados disponíveis.

Descrição dos Dados:

Os dados disponíveis consistem em um conjunto de informações sobre preços de combustíveis coletados em diferentes regiões do Brasil. As colunas incluem:

- Região (sigla)
- Estado (sigla)
- Município
- Revenda

- CNPJ da Revenda
- Nome da Rua
- Número da Rua
- Complemento
- Bairro
- CEP
- Produto
- Data da Coleta
- Valor de Venda
- Valor de Compra
- Unidade de Medida
- Bandeira

Os dados são fornecidos pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) e abrangem um período de tempo significativo, permitindo uma análise histórica dos preços dos combustíveis.

<https://dados.gov.br/dados/conjuntos-dados/serie-historica-de-precos-de-combustiveis-e-de-glp>

Entregáveis:

- Análise Exploratória de Dados:
 - Visualizações e insights sobre os padrões e tendências dos preços dos combustíveis.
 - Identificação de variações regionais, sazonais e relacionadas a fatores específicos.
- Modelo Preditivo de Preços dos Combustíveis:
 - Desenvolvimento de um modelo preditivo utilizando técnicas de aprendizado de máquina para estimar os preços futuros dos combustíveis.
- Relatório Final:
 - Documentação detalhada das análises realizadas, metodologias utilizadas e resultados obtidos.
 - Apresentação dos insights e análises.
- Aplicação Web Interativa:
 - Desenvolvimento de uma aplicação web interativa, usando streamlit, para explorar os dados e visualizar os resultados de forma dinâmica.