

Exercício de apoio - Semana 7: Detecção de anomalias

Use o algoritmo **Fator Local de Anomalia (LOF - Local Outlier Factor)** e o diagrama de caixa para fazer a detecção de anomalias de uma base de dados que contém as negociações de ações da Petrobrás (PETR4) no período entre novembro de 2019 e abril de 2020.

1. Na análise com o LOF, gere um gráfico de dispersão, plotando as anomalias com uma cor diferente dos dados normais. Para a análise, use os atributos **Close** e **Open**. Use o parâmetro **k-vizinhos = 20**. Não precisa definir valor para o parâmetro **contamination**. Dica: insira rótulos nas anomalias para ver suas datas.
2. Em seguida, calcule os limites inferior e superior dos atributos **Close**, **Open** e **Volume** (use o parâmetro **sigma = 1.5**) e gere os diagramas de caixa dos três atributos.

Os dados foram extraídos do Yahoo! Finanças e estão disponíveis nos links abaixo:

Ações da PETR4 - 11/2019 a 04/2020:

<https://br.financas.yahoo.com/quote/PETR4.SA/history?period1=1572566400&period2=1586390400&interval=1d&filter=history&frequency=1d&includeAdjustedClose=true>

Link alternativo:

<https://raw.githubusercontent.com/higoramario/univesp-com360-mineracao-dados/main/petr4-nov-19-abr-20.csv>

Descrição dos atributos:

- Date: data de operação na bolsa.
- Open: valor de abertura.
- High: valor máximo de negociação no dia.
- Low: valor mínimo de negociação no dia.
- Close: valor de fechamento.
- Adj Close: valor de ajuste de fechamento.
- Volume: quantidade de ações negociadas.

O gráfico mostra o valor das ações por dia, mostrando a distribuição da série temporal do período entre 11/2019 e 04/2020.

