

PROGRAMAÇÃO AVANÇADA EM BIG DATA

DESAFIOS PARA AVALIAÇÃO N1

Questão 1: Algoritmo do Boxplot e Detecção de Outliers em Logs de Produção)

Cenário: A equipe de engenharia de dados precisa sanitizar um lote de telemetria contendo 50 registros de tempo de resposta (em milissegundos) capturados por microserviços de um sistema de pagamento digital. A presença de picos de latência (outliers) distorce os modelos preditivos.

Massa de Dados (50 observações):

```
tempos_latencia = [  
    102, 105, 108, 110, 112, 115, 118, 120, 122, 125,  
    128, 130, 132, 135, 138, 140, 142, 145, 148, 150,  
    152, 155, 158, 160, 162, 165, 168, 170, 172, 175,  
    178, 180, 182, 185, 188, 190, 192, 195, 198, 200,  
    205, 210, 215, 220, 225, 480, 520, 610, 750, 890  
]
```

Objetivo: Escreva um script em Python nativo (sem utilizar bibliotecas externas como pandas ou numpy) que:

1. Ordene o conjunto de dados.
2. Calcule o Primeiro Quartil (Q_1 , percentil 25%) e o Terceiro Quartil (Q_3 , percentil 75%) acessando as posições dos índices.
3. Calcule o Intervalo Interquartil ($IQR = Q_3 - Q_1$).
4. Determine o Limite Inferior ($LI = Q_1 - 1,5 \times IQR$) e o Limite Superior ($LS = Q_3 + 1,5 \times IQR$).
5. Percorra o vetor com uma estrutura de repetição (for), filtre e exiba na tela uma lista apenas com os tempos de latência considerados outliers

Questão 2: Ingestão de Logs Não Estruturados e Calculadora Apriori de Produção

Cenário: Um motor de recomendação em tempo real precisa sanitizar descrições textuais brutas vindas de logs de produtos e calcular as métricas de associação a partir de uma base de transações realística do checkout de um varejista.


```
[ "Pão", "Manteiga", "Leite", "Café"],  
[ "Refrigerante", "Snack"],  
[ "Pão", "Manteiga"],  
[ "Pão", "Manteiga", "Café"],  
[ "Cerveja", "Refrigerante"],  
[ "Pão", "Manteiga", "Leite"],  
[ "Pão", "Leite"],  
[ "Pão", "Manteiga", "Snack"],  
[ "Cerveja", "Snack"],  
[ "Pão", "Manteiga", "Leite", "Café"],  
[ "Pão", "Manteiga"],  
[ "Manteiga", "Café"],  
[ "Pão", "Manteiga", "Leite"],  
[ "Cerveja", "Snack", "Manteiga"],  
[ "Pão", "Manteiga", "Café"],  
[ "Pão", "Leite"],  
[ "Pão", "Manteiga", "Leite", "Café"],  
[ "Refrigerante", "Snack"],  
[ "Pão", "Manteiga"],  
[ "Pão", "Manteiga", "Leite"],  
[ "Pão", "Café"]  
]
```

LEVEL UP