Ficamos muito felizes com o seu interesse nesta vaga e esse teste é de grande importância para que possamos avaliar o seu conhecimento técnico, criatividade e organização no desenvolvimento da tarefa. Nosso trabalho diário é desenvolver e implementar soluções de dados que aumentam a performance dos nossos clientes nos principais sites de varejo online. Por isso, criamos este desafio técnico para simular uma atividade que é realizada pelos nossos times especialistas. Nessa tarefa, você deverá encontrar os 10 produtos com a maior variação de preço, entre o valor encontrado no varejo online (Americanas) e o estipulado pelo fabricante, e os 10 produtos que apresentam maior indisponibilidade. Nosso cliente deseja observar com essa task se existem produtos que estão sendo vendidos com valores muito distantes do ideal. Para isso, disponibilizamos dados coletados nos varejos em diferentes datas para que possa desenvolver essa atividade. Alguns pontos importantes sobre a estrutura dos nossos dados:

* "retailerPrice" -- é preço encontrado no varejo;
* "manufacturerPrice" -- é preço indicado pelo fabricante;
* "priceVariation" -- é a variação do preço, pode ser negativa ou positiva.
* “available” – se o produto está disponível para compra

O escopo do teste deve incluir, mas não se restringindo:

* Ingestão dos dados;
* Armazenamento dos arquivos;
* Tratamento dos dados;
* Cálculo das métricas;

Desejáveis:

* Aplicação de PyLint ou similar;
* Aplicação de Clean Code;
* Diagrama da pipeline;
* Readme;
* Coverage e Testes unitários;

Os arquivos para o teste estão em formato .zip no link;

<https://drive.google.com/file/d/14nMMC19vpQ8s69TqCkQZh7SxJxlpaOzW/view?usp=sharing>

Esperamos que o resultado seja entregue em um repositório GIT.

Boa sorte!