



Seja bem-vindo(a) à tarefa de casa de Dados da Hype!

Para a correta avaliação, consideramos uma empresa fictícia chamada Acme Inc.

Sua tarefa será realizar algumas análises para ajudar a Acme a reduzir seus custos de operação.

Temos 3 questões a seguir para serem resolvidas até a data limite, que abordam parte das tarefas do dia-a-dia do departamento de Dados da empresa.

Regras para resolução:

- Você pode usar a ferramenta de modelagem de sua escolha para resolver a questão 1
- Você pode utilizar a linguagem de programação de sua escolha para resolver as questões 1, 2 e 3
- Para as questões 2 e 3, você deve primeiro realizar a modelagem dos dados em um sistema de banco de dados de sua escolha (SQL ou NoSQL) para, em seguida, utilizar sua linguagem escolhida para buscar e avaliar os dados que precisa. É possível responder às questões usando a própria linguagem do banco caso prefira.

Será avaliado:

- A forma com que você estrutura seu projeto;
- O uso de boas práticas para a resolução das questões;
- Os resultados obtidos;
- A apresentação dos resultados;
- A forma com que você comunica os pontos acima;

Não será avaliado:

- Linguagem de programação utilizada (você pode usar a que achar melhor)
- Banco de dados escolhido
- Tempo para resolver, mas ainda limitado ao tempo estipulado para a entrega do exercício

Apresentação dos resultados:

- Monte uma apresentação com os resultados de cada questão, que deve ser enviada para nós até a data de entrega do exercício.
- Não pedimos o código fonte mas podemos pedir para você nos mostrar sua implementação na data da entrevista, portanto, tenha o código disponível nesta data.



1 - A Acme Inc, por ser uma empresa brasileira que lida com clientes pessoas físicas, utiliza diversos indicadores de mercado, dentre eles, o IPCA, que reflete no poder de compra da população. Esse dado é fornecido de forma aberta pelo IBGE no [seguinte link \(JSON\)](#). Para efeitos do exercício, este dado já está descontando a sazonalidade.

Você deve:

- A) Utilizar a linguagem de programação de sua escolha para extrair esse dado para um banco de dados também de sua escolha (SQL ou NoSQL)
- B) Utilizar os dados previamente salvos no banco de dados para mostrar, utilizando diretamente a linguagem nativa do banco de dados de sua escolha:
 - O **ano** com maior estabilidade, ou seja, menor variação do indicador
 - O **mês** com maior aumento no indicador
- C) Uma explicação de até 5 linhas do motivo do aumento observado no período em torno do mês com maior aumento no indicador.

2 - A Acme Inc. utiliza um serviço terceirizado de computação distribuída para a execução de diversas funções (código que executa alguma tarefa específica). O modelo de precificação deste fornecedor ocorre de acordo com o tempo que cada função permanece rodando dentro da plataforma.

A Acme é cobrada 0.000015 R\$ por milissegundo de uso de cada invocação de qualquer uma de suas funções nesta infraestrutura, sendo o número de milissegundos sempre arredondado para cima a cada centena, desta forma, uma invocação que leva 33 milissegundos seria cobrada pelo tempo de 100 milissegundos.

Este modelo sofreu uma alteração em 13/06/2020 de forma que não existe mais esse arredondamento e o tempo de execução é sempre igual ao tempo que será cobrado da Acme.

No arquivo **executions.csv** temos o log de cada execução de cada função da Acme e o tempo em milissegundos que essa execução levou para ser concluída. No arquivo **functions.csv**, temos os dados das funções e o tempo médio relativo à fatores externos que influenciam em sua execução.

Sua tarefa é:

- A) Montar o diagrama ER que descreve o modelo de dados dos dois arquivos



- B) Calcular o custo total que a Acme teve levando em conta toda a base de dados de execuções
- C) Calcular quanto a Acme deixou de gastar com a implementação do novo modelo de precificação
- D) Calcular qual seria o custo total se todas as transações estivessem no modelo antigo
- E) Calcular qual seria o custo total se todas as transações estivessem no modelo novo
- F) Mostrar um gráfico comparativo com os cenários descritos em B, C e D ao longo do tempo.

3 - Para efeitos de redução de custo, a Acme precisa identificar quais funções (dos arquivos em anexo) podem ser otimizadas pelo departamento de Desenvolvimento. Consideramos uma função não otimizada aquela que possui alta variação em tempo de execução em relação à média dos tempos de execução, desde que ela não acesse algum sistema externo. Se ela acessa algum sistema externo, deve-se descontar do cálculo a variabilidade externa média desta função (já inclusa nos dados em anexo). Desta forma, sua tarefa será:

- A) Identificar quais funções precisam ser revisadas. Considere 30% como o limite de tolerância para a variação em relação à média e que todas as funções que estiverem além deste limite precisam ser revisadas.
- B) Mostrar um gráfico que apresenta a variação ao longo do tempo, a média, os limites de tolerância e os pontos nos meses em que a variação passa do limite de forma diferenciada.