



Desafio para a vaga de Data Analyst II

Olá, parabéns por ter chegado a essa nova fase. Nós elaboramos um desafio bem bacana para conhecermos melhor a sua forma de trabalho.

Pelo fato de ser uma grande empresa do ramo do varejo, a Bemol está sempre desenvolvendo indicadores, metas e analisando dados de vendas. Alguns indicadores mais comuns contêm informações sobre margem e volume de vendas, faturamento, produtos e categorias mais vendidas, varejo e farma, análise de séries temporais, entre outros. Por isso, a diretoria financeira deseja estudar os resultados comerciais da empresa, observar tendências e projetar como será o desempenho para o ano seguinte, com o objetivo de traçar as próximas estratégias. Porém, precisamos da sua ajuda para enxergar mais oportunidades e gerar novas ideias.

Ao final deste documento, na seção de Observações, você encontrará um link para acessar o drive onde você encontrará os seguintes arquivos:

- Arquivo `base_vendas_desafio.csv`: base de dados necessária para o desafio;
- Arquivo `dicionario.txt`: dicionário de dados para melhor compreensão das informações;
- Cópia do arquivo do desafio em pdf

Acesse os arquivos disponibilizados no drive pelo link a seguir:

Você deve gerar insights que sejam relevantes para entender o desempenho comercial da Bemol nos últimos anos e comunicar de maneira clara e eficaz suas descobertas obtidas com a análise dos dados disponibilizados. Dois indicadores fundamentais para a diretoria são: a margem de lucro financeira (em R\$) e a margem de lucro percentual (em %). Além destes dois, é importante buscar trazer outros indicadores, além de apontar que perguntas de negócio poderiam ser respondidas com base nos entendimentos que você obteve. Para o cálculo de margem considere o preço de venda e o preço de custo total de cada produto vendido. Realize todos os processos de limpeza/tratamento que forem necessários e, ao identificar possíveis dados com valores incorretos, retire estes registros da base de dados antes de prosseguir com as análises. Considere que as colunas de tipo numérico com valores nulos devem ser convertidas ao valor 0 (zero). Faça uso de suas habilidades em **programação, estatística e visualização de dados**, com algumas das tecnologias mais comuns na área de data science.

Além disso, como **análise complementar e opcional**, você pode realizar uma projeção futura relacionada a algum indicador de vendas, utilizando de métodos estatísticos, modelo de aprendizado de máquina, ou outro de sua preferência. Não é obrigatório, mas seria desejável adicionar o projeto, bem como suas modificações, em um repositório git.

Não temos restrição quanto à *stack*, mas gostaríamos de algo como: **Python, SQL, Spark**. Sinta-se livre para usar sua **stack preferida**. Como ocorre na maioria dos problemas em *data science*, algumas explicações sobre a solução podem ser requeridas durante o processo de seleção. Não esqueça de nos contar sobre seu entendimento dos dados, a modelagem utilizada, as métricas escolhidas e análises realizadas, nos mostrando suas principais

características. É importante que você deixe o mais claro possível suas suposições e decisões de projeto feitas ao longo do desafio.

O que é fundamental no seu trabalho:

- Arquivo com o código (na linguagem de sua preferência) contendo a análise exploratória dos dados com informações claras dos métodos e técnicas utilizadas;
- Análise estatística com insights trazidos do dataset e gerada visualmente com bibliotecas de DataViz;
- Apresentação gráfica dos resultados utilizando ferramentas de BI (Power BI, Data Studio, Tableau, etc);
- Apresentação em poucos slides com informações objetivas sobre o resultado;
- Criar um glossário de dados do desafio, descrevendo informações como tabelas e colunas novas que foram criadas, listagem e descrição de funções desenvolvidas, metodologias, entre outros. O formato de arquivo e estrutura do glossário ficam a seu critério;
- Entrega de qualquer artefato adicional utilizado (planilhas, uso de softwares, arquiteturas, etc);

Observações:

- O conteúdo dos arquivos compartilhados (o texto deste desafio, a base de dados e o dicionário de dados) não pode ser copiado/divulgado;
- Você enviará os arquivos-resposta com o título: Desafio: Bemol Digital Data Analyst II - Seu Nome para talent@bemol.com.br;
- [Base para o desafio \(clique aqui\)](#)

Tenha um ótimo desafio e divirta-se! :)
Qualquer dúvida estamos à disposição.

Time de Data Science