Teste para Estagiário de Analytics Quod

## Instruções Gerais

### Ferramentas Permitidas:

* Você pode usar recursos online como documentações, tutoriais e fóruns de programação e ciência de dados.
* **Uso de ferramentas de IA**: Seu uso é permitido, porém vale observar que a prova tem caráter conceitual, ou seja, será avaliado a artir de entrevistas o conhecimento aplicado à prova e suas variantes. Por esse motivo, sugerimos utilizar o máximo possível do seu conhecimento pessoal.

### Entrega:

* Crie um repositório público no GitHub, GitLab ou Bitbucket com o nome Teste\_Analytics\_NomeSobrenome.
* Faça upload de todos os seus códigos, arquivos de dados simulados e um documento (em formato Word, Markdown ou outro) com respostas textuais e explicações.
* Inclua um arquivo README.md explicando a estrutura do repositório, como executar os scripts e quaisquer dependências necessárias.
* Envie um e-mail para [o time de RH da Quod] com o link do seu repositório.
* O e-mail deve ter o assunto "Entrega do Teste para Estagiário de Analytics - [Nome e Sobrenome]".

**Prazo:** O teste deve ser entregue em até 9 dias a partir da data de recebimento.

**Observação:** Documente quaisquer suposições feitas durante a simulação dos dados ou análises no README.md.

## Tarefas

### Parte 1: Programação em R ou Python

Escolha entre R ou Python para resolver os problemas a seguir. Inclua os scripts com comentários explicativos e, quando necessário, forneça explicações adicionais no documento.

#### Limpeza e Análise de Dados de Vendas

* + Crie um script para simular um dataset de vendas com pelo menos 50 registros, contendo as colunas: ID, Data, Produto, Categoria, Quantidade, Preço. O período dos dados deve ser de 01/01/2023 a 31/12/2023.
  + Realize a limpeza dos dados, incluindo:
    - Tratamento de valores faltantes
    - Remoção de duplicatas
    - Conversão de tipos de dados, se necessário
  + Salve o dataset limpo em um arquivo data\_clean.csv.
  + Utilizando o dataset limpo, realize as seguintes análises:
    - Calcule o total de vendas (Quantidade \* Preço) por produto.
    - Identifique o produto com o maior número de vendas totais.
  + **Salve seu código e nomeie adequadamente.**

#### Análise Exploratória de Dados de Vendas

* + Com o mesmo dataset simulado da tarefa anterior, crie visualizações/gráficos e com eles:
    - Crie um gráfico de linha mostrando a tendência de vendas ao longo do tempo, analisando os dados mensalmente.
    - Identifique e descreva pelo menos dois padrões ou insights interessantes que você observou nos dados.
  + **Salve seu código e nomeie adequadamente.**

### Parte 2: SQL

Com base nas tabelas geradas no exercício anterior, escreva consultas SQL para:

* Listar o nome do produto, categoria e a soma total de vendas (Quantidade \* Preço) para cada produto. Ordene o resultado pelo valor total de vendas em ordem decrescente.
* Identificar os produtos que venderam menos no mês de junho de 2024.

**Salve suas consultas e a explicação da lógica em um arquivo chamado consultas\_sql.sql.**

### Parte 3: Interpretação de Resultados

* Com base nas análises realizadas nas Partes 1 e 2, escreva um breve relatório (máximo de 300 palavras) resumindo os principais insights encontrados e sugerindo possíveis ações baseadas nesses insights.

**Salve este relatório como relatorio\_insights.md ou relatorio\_insights.pdf.**

**Finalização da entrega:** Para fluidez da avaliação oral é esperado que o candidato faça uma apresentação da atividade realizada, bem como os resultados obtidos. Tal apresentação pode ser realizada no software de sua preferência.

#### **Critérios de Avaliação**

* Clareza e organização dos arquivos, do código e da documentação
* Interpretação adequada dos resultados
* Uso apropriado de técnicas estatísticas e de análise de dados
* Qualidade e relevância das visualizações criadas
* Criatividade na abordagem dos problemas
* Capacidade de simular dados relevantes para os problemas propostos
* Habilidade em extrair insights significativos dos dados

**Boa sorte!**