

Case Ton | Análise de dados

Neste estudo de caso, temos alguns cenários hipotéticos com o objetivo de avaliar a sua forma de resolver problemas.

Para a resolução das questões a seguir, foram disponibilizadas as tabelas de dados necessárias em formato csv da forma descrita no anexo I.

O estudo deve ser realizado em **Python**, usando **SQL** (qualquer sintaxe) para a questão 4.

Entregáveis

- Documento word/pdf contendo as respostas das questões 1 a 5, com limite de 8 páginas.

Sugestões para ter no documento:

- Tabelas e/ou gráficos que justifiquem as respostas;
- Texto **breve** explicando o raciocínio e motivos de tomada de decisão em cada questão;

Contexto

Você é o responsável por dados e análises no time de Relacionamento com o Cliente. Este time é responsável por cuidar dos nossos clientes, tirando todas as dúvidas e os ajudando a fazer melhor uso dos nossos produtos.

Nos últimos tempos tivemos alguns problemas e precisamos da sua ajuda para tomarmos as melhores decisões para melhor atender nossos clientes.

Questões

1. Foi implementada no dia 1 de agosto uma funcionalidade nova no site do Ton que acredita que **irá diminuir a quantidade de chamados recebidos. Avalie se a implementação teve um impacto estatisticamente relevante.**
2. Foi notado pelo nosso time que **clientes recém credenciados possuem um comportamento específico em relação à quantidade de chamados realizados logo após seu credenciamento.** Identifique e prepare alguma forma de **visualizar** essa diferença notada pelos agentes.
3. Para conseguir melhorar a qualidade do nosso atendimento ao cliente, foi proposto que a **alguns chamados tenham prioridade de atendimento.** Levando em consideração o motivo dos chamados, **como você estabeleceria essa prioridade da fila de espera? Sugestão:** Proponha uma clusterização da base de clientes que julgue atender da melhor forma a necessidade apresentada.
4. Escreva duas queries em **SQL** que retornem:
 - i) o volume de chamados por semana dos últimos três meses para cada cluster de clientes proposto na questão 3.
 - ii) uma série histórica dia a dia, que para cada dia retorne o número de chamados referentes aos últimos 30 dias (por exemplo, para o dia 30/01, a query deve retornar a quantidade de chamados totais entre os dias 01/01 e 30/01)

5. Questão de estudo livre: Como é esperado de um analista de dados, traga insights relevantes sobre pontos de atenção e possíveis melhorias que você identificou ao analisar ambas as bases de dados. Além disso faça pelo menos duas perguntas relacionadas aos dados que julgue relevantes para análise, você também pode sugerir outros dados para enriquecer os aqui disponibilizados.

ANEXO I

Tabela 'cases': Chamados recebidos pelo time de Relacionamento com o Cliente.

Coluna	Descrição
accountid	Chave primária da tabela
date_ref	Data do chamado
channelid	Canal do chamado (chat = 1; telefone = 2; e-mail = 3)
waitingtime	Tempo de espera para atendimento (segundos)
missed	Flag de chamado atendido
pesquisa_de_satisfa_o_c	Flag de pesquisa de satisfação preenchida
assunto	Assunto do chamado
Id	Chave da tabela

Tabela 'creds': Clientes credenciados.

Coluna	Descrição
cred_date	Data de credenciamento
shipping_address_city	Endereço de entrega – cidade
shipping_address_state	Endereço de entrega – estado
max_machine	Produto adquirido
accountid	Chave secundária da tabela