

SQL para análise de dados



PERGUNTAS FREQUENTES



SQL para análise de dados



Dica: para encontrar rapidamente a palavra que procura aperte o comando CTRL+F e digite o termo que deseja achar.

- Extensão ipynb
- Google Colab
- Tabelas em SQL
- Restrições em colunas
- Chave primária
- SELECT DISTINCT
- Selecionar colunas
- Ordenação
- Tabela no Athenas
- Coluna no Athenas
- Id_transacao no MySQL
- Operador IN
- Operador LIKE
- Wildcards
- Where com like
- UNION
- Subqueries
- Subquery e a cláusula FROM
- Subquery correlacionada e não correlacionada



Extensão ipynb

O que são arquivos com extensão ipynb?

Refere-se a um arquivo no formato do Jupyter Notebook, uma aplicação web de código aberto que permite criar e compartilhar documentos interativos contendo código, visualizações, texto e outros elementos. Os arquivos .ipynb são usados para criar, salvar e abrir notebooks no Jupyter e também no Google Colab.



Google Colab

O que é Google Colab?

Também conhecido como Google Colaboratory, é uma plataforma que suporta a leitura e execução de arquivos com a extensão ".ipynb". Ele é baseado no ambiente Jupyter Notebook e permite que você abra, edite, execute e crie notebooks diretamente em seu navegador.



Tabelas em SQL

Como faço para remover duplicidades numa tabela?

Caso tenha registros duplicados em uma coluna, você pode, em conjunto com a cláusula "SELECT", usar o comando "DISTINCT" para remover as duplicidades.



Restrições em colunas

É possível definir restrições ou regras para as colunas da tabela?

Sim, é possível definir restrições e regras para as colunas da tabela. Alguns exemplos incluem restrições de NOT NULL (para garantir que a coluna não seja nula), restrições de UNIQUE (para garantir valores únicos) e restrições de FOREIGN KEY (para estabelecer relacionamentos com outras tabelas).



Chave primária

Como definir uma chave primária ao criar uma tabela?

Para definir uma chave primária, você pode usar a cláusula PRIMARY KEY após a definição da coluna desejada.



SELECT DISTINCT

O comando SELECT DISTINCT sempre traz um campo só?

Não, esse comando não está limitado a trazer apenas um campo, ele pode ser usado para obter valores únicos de um campo específico ou para obter combinações únicas de valores em várias colunas, dependendo das colunas que você especificar após o comando.



Selecionar colunas

Como selecionar todas as colunas de uma tabela?

Para selecionar todas as colunas de uma tabela, você pode usar o asterisco (*) como curinga.



Ordenação

Como ordenar os resultados em ordem crescente ou decrescente?

A cláusula ORDER BY é usada para ordenar os resultados de uma consulta. Por padrão, a ordenação é feita em ordem crescente, mas você pode especificar a ordem decrescente, se necessário. É possível ordenar por múltiplas colunas, basta especificar na cláusula ORDER BY, separadas por vírgula. A ordenação ocorre na ordem especificada.



Tabela no Athenas

Por que quando adicionamos uma coluna com o ADD COLUMNS no Athena ele altera os DATATYPES de todo o resto da coluna?

A estrutura da tabela no Athena é imutável, ou seja, quando uma nova coluna é adicionada, ela cria uma versão da tabela com a nova coluna e os tipos de dados atualizados para todas as colunas existentes. Essa abordagem é necessária para garantir a consistência e a integridade dos dados armazenados na tabela. Portanto, é recomendado revisar e ajustar as consultas existentes para se adequarem ao novo esquema da tabela, a fim de otimizar o desempenho das consultas no Athena.



Coluna no Athena

É possível deletar uma coluna no Athena?

Por enquanto não é possível. Após a criação da tabela, o que pode ser feito é utilizar o REPLACE COLUMNS, substituindo um novo nome e tipo do campo da tabela.



Id_transacao no MySQL

É possível gerar automaticamente o id_transacao no MySQL?

Sim, basta utilizar a função `AUTO_INCREMENT` na criação da tabela, nesse caso ele vai gerar uma ID, que é categorizado e gerado automaticamente, assim aumentando o valor do ID de um a um.



Operador IN

Só é possível usar o operador IN com strings?

Não, o operador IN em SQL não se limita apenas a strings. Ele pode ser usado com diversos tipos de dados, como strings, números, datas e outros tipos. Ele é usado para verificar se um valor está presente em uma lista de valores especificada. Essa lista pode conter valores de diferentes tipos de dados, desde que sejam compatíveis com o tipo de dados da coluna em que a comparação está sendo feita.



Operador Like

Como usar o operador LIKE para realizar buscas por padrões?

O operador LIKE é usado em conjunto com caracteres curinga (%) e (_) para buscar padrões em uma coluna.



Wildcards

Os wildcards estão relacionados à padronização das nomenclaturas em um banco de dados?

Os wildcards são usados para fazer correspondência parcial ou padrões em consultas de pesquisa, filtragem ou seleção de dados. Eles não estão diretamente relacionados à padronização das nomenclaturas em um banco de dados, mas são úteis para buscar dados com base em padrões específicos.



Where com like

A acentuação interfere na busca no filtro where com like em SQL?

Sim, a acentuação pode interferir na busca no filtro WHERE com LIKE em consultas SQL, dependendo da configuração do banco de dados e do tipo de collation (ordenação) usado. Caso você precise realizar buscas considerando ou ignorando a acentuação de forma específica, é recomendado ajustar a collation da coluna ou utilizar funções específicas para tratar a acentuação, como COLLATE ou funções de substituição de caracteres, dependendo do banco de dados utilizado.



UNION

Para fazer a operação UNION as colunas precisam ter o mesmo nome?

Não, o UNION combina os resultados de duas ou mais consultas em um único conjunto de resultados, independentemente dos nomes das colunas. Ao usar essa operação, é importante garantir que o número e os tipos de colunas correspondam entre as consultas. As colunas nas diferentes consultas devem estar na mesma ordem e ter tipos de dados compatíveis. As colunas serão identificadas pelos índices de posição no resultado, não pelos nomes.



Subqueries

As subqueries podem retornar múltiplas colunas?

Sim, as subqueries podem retornar múltiplas colunas, assim como as consultas regulares. Nesse caso, é importante garantir que o número de colunas retornado pela subquery seja compatível com a cláusula que a utiliza.



Subquery e a cláusula FROM

É possível usar uma subquery em uma cláusula FROM?

Sim, é possível usar uma subquery na cláusula FROM para criar uma tabela temporária que será usada na consulta externa. Isso permite a geração dinâmica de dados a partir de outras tabelas ou subqueries.



Subquery correlacionada e não correlacionada

Qual é a diferença entre uma subquery correlacionada e uma subquery não correlacionada?

Uma subquery não correlacionada é independente da consulta externa, ou seja, ela pode ser executada isoladamente. Já uma subquery correlacionada faz referência a uma coluna da consulta externa, e seu resultado depende do contexto da consulta externa.



Bons estudos!

