

SQL

Criação de banco de dados

Você sabia que seu material didático é interativo e multimídia? Isso significa que você pode interagir com o conteúdo de diversas formas, a qualquer hora e lugar. Na versão impressa, porém, alguns conteúdos interativos ficam desabilitados. Por essa razão, fique atento: sempre que possível, opte pela versão digital. Bons estudos!

Na criação de um banco de dados é necessário realizar o planejamento do mesmo, decidir as tabelas necessárias, criar algumas regras de convenção, definir as propriedades do banco de dados e questões relacionadas sua internacionalização. As propriedades de um banco de dados são de vital importância para determinar como os dados serão armazenados.

Planejamento de um banco de dados

O planejamento de um banco de dados não trata somente na forma como os dados serão armazenados, mas também em como sua estrutura é definida. Para isso temos cinco passos principais:



Fonte: Shutterstock.

Coletar informações

Deve-se ter um entendimento do trabalho ao qual o seu banco de dados se destina. Ele deverá fornecer todas as informações necessárias para alcançar seu objetivo. É importante entender qual a abrangência de seu uso e público-alvo, pois muitas informações deverão ser úteis àqueles que no futuro utilizarão seu banco de dados. Também é necessário identificar limitações e problemas, para que na definição de sua estrutura você possa ultrapassá-los.

Identificar suas principais estruturas

Identificar suas principais entidades, ou seja, suas tabelas, pois elas serão gerenciadas pelo banco de dados. Estas tabelas representarão algo tangível ou intangível, e isso quer dizer que podemos representar pessoas, locais, clientes, cidades e também uma venda, uma reclamação, uma autorização de compra, um período de pagamentos de impostos, entre outras coisas.

Modelar a estrutura

A parte da modelagem de estrutura pressupõe que as principais tabelas estejam identificadas, então faz-se necessário gerar uma documentação, que é o Diagrama de Entidades e Relacionamentos (DER). Por se tratar de um registro intelectual de criação, este documento servirá para que outras pessoas entendam como foi idealizado o modelo de abstração para a resolução de um problema, e por isso deve estar sempre atualizado.

Identificar nas estruturas os tipos de dados

Deve-se identificar detalhadamente os tipos de dados que devem ser armazenados, números inteiros ou decimais, cadeia de caracteres, datas, etc. A classificação destas informações também se dá em relação ao seu conteúdo:

- Dados brutos (nomes, endereços telefones, nascimento)
- Dados de categorização (verdadeiro/falso, certo/errado, casado/solteiro)
- Dados de identificação (identificador único de um registro na tabela, não pode ser nulo, ou repetir-se, ex. cliente_id, produto_id)
- Dados de relação ou referência (armazenam a relação entre duas tabelas, é através dela que o banco de dados gerencia integridades, consultas e relacionamentos).

Identificar quais são os seus relacionamentos

Razão de um banco de dados relacional existir: relacionar dados, associar informações sobre várias tabelas de seu banco de dados. Pode-se fazer as mais variadas combinações, pois os mecanismos do banco de dados são desenvolvidos para isso, desenvolver relações com definições lógicas de agregação ou desagregação de dados.

Internacionalização de um banco de dados

Quando desenvolve-se um banco de dados é necessário prestar atenção na sua abrangência de utilização, pois os dados que serão armazenados, devem respeitar as regras de escrita e gramática ou representação de cada país. Os sistemas gerenciadores de bancos de dados têm algumas cláusulas especiais para lidar com estas questões. Elas são CHARSET e COLLATION.



CHARSET

Conjunto de caracteres

A cláusula CHARSET designa um conjunto de símbolos e codificações e, como eles são representados binariamente.



COLLATION

Agrupamentos

A cláusula COLLATION é o conjunto de regras para a comparação de caracteres num conjunto de caracteres.

SAIBA MAIS

Você já pode ter acessado algum aplicativo ou site na internet que exibia algumas palavras com a grafia incorreta. Talvez não seja um erro de conteúdo, mas uma falha de apresentação deste conteúdo. Por isso a necessidade de entendimento da utilização do CHARSET e COLLATION.

CHARSET

Há muitos alfabetos inteiros, às vezes múltiplos alfabetos ou sistemas de escrita orientais com milhares de caracteres, juntamente com muitos símbolos especiais e sinais de pontuação. O MySQL suporta inúmeros conjuntos de caracteres, e seus respectivos agrupamentos de regras de comparação.



A B C D E F G H I J K L M N O P Q R
S T U V W X Y Z 0 1 2 3 4 5 6 7 8 9
* / > < & # N º ? ! @ % = + - \$ € £ (. ,)

Para exibir todos os conjuntos de caracteres instalados utilize a instrução:

```
SHOW CHARACTER SET;
```

Vamos entender os dois conjuntos de caracteres, o “**latin1**” e o “**utf8**”.

```
SHOW CHARACTER SET WHERE charset LIKE 'latin1'
```

Tem especificamente em sua descrição (DESCRIPTION) “cp 1252 West European” e, como padrão está associado o conjunto de regras (DEFAULT COLLATION) com o conteúdo “latin1_swedish_ci”.

```
SHOW CHARACTER SET WHERE charset LIKE 'utf8';
```

Tem especificamente em sua descrição (DESCRIPTION) “UTF-8 Unicode” e, como padrão está associado o conjunto de regras (DEFAULT COLLATION) com o conteúdo.

Diferença entre as instruções

O “**utf8**” representa muitas línguas e o “**latin1**” uma parte destas línguas. Por esta razão aplicações que terão sua utilização com abrangência internacional, deve-se utilizar o padrão UTF-8.

COLLATION

Toda a cadeia de caracteres (CHARSET) apresenta um conjunto de regras associado (COLLATION), mas para exibir todas as instruções instaladas deve-se utilizar a instrução:

```
SHOW COLLATION;
```

Para visualizar o conjunto de regras do “utf8”, utilize a instrução:

```
SHOW COLLATION WHERE collation LIKE 'utf8_general_ci';
```

Alguns conjuntos de regras têm identificado ao final um “_ci”, ou “_cs”.



As letras “CI”

case insensitive

Não há distinção entre maiúsculas e minúsculas. Pesquisas feitas utilizando a palavra “moura”, retornarão “Moura”, “MOURA” ou “moura”.



As letras “CS”

case sensitive

Há distinção entre maiúsculas e minúsculas. Pesquisas feitas utilizando a palavra “moura”, retornará apenas “moura”.

Linguagem de Definição de Dados (DDL – Data Definition Language)

Criação de um banco de dados

Para criar um banco de dados deve-se utilizar as instruções da classe da linguagem de definição de dados (DDL).

A instrução a seguir cria um banco de dados com o nome especificado:

```
CREATE {DATABASE | SCHEMA} [IF NOT EXISTS] db_name  
[create_specification] ...  
create_specification:  
[DEFAULT] CHARACTER SET [=] charset_name  
| [DEFAULT] COLLATE [=] collation_name
```

Saiba Mais

As características do banco de dados são armazenadas no arquivo **db.opt** no diretório do banco de dados.

A seguir, veja as instruções para criação do banco de dados chamado “mundo”

Criar banco de dados

Para criar o banco de dados “mundo”, com o CHARSET “utf-8” e COLLATION “Utf8_general_ci”, a instrução será a seguinte:

```
CREATE DATABASE mundo  
  DEFAULT CHARSET = utf8  
  DEFAULT COLLATE = utf8_general_ci;
```

O banco de dados foi criado, e se você utilizar a instrução `SHOW DATABASES` é possível vê-lo.

Alerta de que o banco já existe

Se for utilizada a instrução de criação de banco de dados e o mesmo já existir, será visualizada uma mensagem de erro. Para que isso não ocorra deve-se utilizar a instrução da seguinte maneira:

```
CREATE DATABASE IF NOT EXISTS mundo  
  DEFAULT CHARSET = utf8  
  DEFAULT COLLATE = utf8_general_ci;
```

Agora não terá um erro, apenas um alerta de que o banco já existe.

Alterar CHARSET ou COLLATE

Para alterar o CHARSET ou COLLATE do banco de dados, deve-se usar a instrução:

```
ALTER {DATABASE | SCHEMA} [db_name]  
  alter_specification ...  
  alter_specification:  
  [DEFAULT] CHARACTER SET [=] charset_name  
  | [DEFAULT] COLLATE [=] collation_name
```

Para alterar o CHARSET do banco de dados mundo para “latin1”:

```
ALTER DATABASE mundo CHARSET = latin1;
```

Apagar o banco de dados

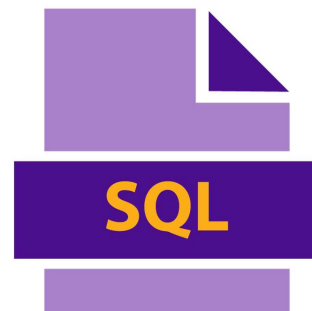
Para apagar o banco de dados, deve-se usar a seguinte instrução:

```
DROP {DATABASE | SCHEMA} [IF EXISTS] db_name
```

A execução deste comando irá apagar o banco de dados especificado em `db_name` e a cláusula `IF EXISTS` previne que não será gerado um erro se a base não existir.

Script

Script (roteiros de instruções SQL) é um arquivo texto onde estão todas as instruções SQL. A criação de scripts é bastante utilizada para criação de bancos de dados. Em um script pode-se adicionar quaisquer comandos SQL, porém eles devem estar numa ordem lógica de execução.



Fonte: Shutterstock.

Manter uma documentação atualizada sobre todas as definições e alterações num repositório de banco de dados é vital para a manutenção de scripts na sustentação de uma área de desenvolvimento.