

Bancos de Dados

Aula 02 – Conceitos Básicos de Bancos de Dados

Resultados esperados

- Compreender o que são bancos de dados e porque eles são utilizados.
- Visualizar aplicações de bancos de dados no dia-a-dia.
- Ter um vislumbre de porque vamos utilizar bancos de dados relacionais baseados em SQL em detrimento de usar um arquivo texto ou um arquivo excel.
- Diferenciar bancos de dados e sistemas gerenciadores de bancos de dados (SGBD).
- Entender o que é SQL.
- Explicar que vamos focar em bancos de dados relacionais e que neste tipo de banco de dados, dados são armazenados em tabelas mais ou menos como uma tabela do Excel com uma linha de cabeçalho.
- Definir tabela, schema e linhas (tuplas).
- Entender e praticar como criar uma tabela, verificar se ela foi criada corretamente, inserir dados e listar dados usando SQL. Compreender como uma tabela em SQL pode ser representada usando um diagrama.
 - Comandos que serão estudados:
 - Criar tabela: comando CREATE TABLE.
 - Verificar se a tabela foi criada no sqlite:
SELECT name, sql FROM sqlite_schema WHERE type = 'table';
 - Inserir dados: INSERT
 - Listar dados inseridos: SELECT simples.
 - Fazer um exercício prático.
- Mostrar como executar estes comandos em Python usando a biblioteca sqlite3.

Motivação

Os programas estudados na primeira disciplina interagiam com o usuário usando a tela e o teclado, armazenavam dados na memória do computador e usavam o processador para transformar dados usando operações matemáticas e lógicas.

Nestes programas, os dados do programa só existem enquanto o programa estiver em execução.

Alguns programas precisam armazenar dados por períodos longos e mesmo se o computador for desligado, e.g., universidade, loja, redes sociais, etc.

Memórias secundárias como HD e SSD podem ser usados para armazenar dados de forma persistente (não-volátil).

Tipos de Aplicações que usam Bancos de Dados

Aplicações de BD Tradicionais: Por exemplo, quando vamos ao banco para depositar ou retirar fundos, fazemos uma reserva de hotel ou de voo, acessamos o catálogo de uma biblioteca virtual para procurar uma referência bibliográfica, ou compramos algo on-line — como um livro, um brinquedo ou um computador —, provavelmente essas atividades envolverão alguém ou algum programa de computador que acessa um banco de dados.

Aplicações de BD Multimídia: sites de streaming de músicas e vídeos.

Os sistemas de informações geográficas (GIS — Geographic Information Systems) podem armazenar e analisar mapas, dados sobre o clima e imagens de satélite.

Sistemas de data warehousing e de processamento analítico on-line (OLAP — On-Line Analytical Processing) são usados em muitas empresas para extrair e analisar informações comerciais úteis de bancos de dados muito grandes, para ajudar na tomada de decisão.

Definições Básicas

Sistema gerenciador de bancos de dados (SGBD): É uma coleção de dados interrelacionados e um conjunto de programas para acessar a manipular estes dados. Os mais usados são Microsoft SQL Server, Oracle, MySQL e PostgreSQL. SQLite também é muito usado em projetos menores pela simplicidade.

Definição Alternativa: Um sistema gerenciador de banco de dados (SGBD — Database Management System) é uma coleção de programas que permite aos usuários criar e manter um banco de dados. O SGBD é um sistema de software de uso geral que facilita o processo de definição, construção, manipulação e compartilhamento de bancos de dados entre diversos usuários e aplicações. Observação: O termo sistemas de bancos de dados é usado em [1] para se referir a SGBDs, enquanto em [2] é usado para a união do banco de dados com o software de SGBD. Na nossa disciplina, o sqlite3 fará o papel de SGBD e depois usaremos o PostgreSQL.

Banco de dados é o nome que se dá a esta coleção de dados. Em geral, o banco de dados é armazenado em um conjunto de arquivos em uma memória secundária (HD, SSD, etc.). Em algumas aplicações excepcionais podem ser criados bancos de dados em memória primária.

Um banco de dados representa algum aspecto (subconjunto) do mundo real, chamado de **minimundo** ou de **universo de discurso** (UoD — Universe of Discourse). O minimundo contém os dados que precisamos armazenamos pelo sistema para alcançar um objetivo organizacional.

Tarefas do SGBD

- Fornecer uma maneira de inserir, atualizar, deletar e recuperar informações do banco de dados que seja conveniente, eficiente. Arquivos simples e tabelas do excel em geral são menos eficientes que SGBDs.
- Além disso, o sistema de bancos de dados deve garantir a segurança das informações armazenadas, apesar de falhas do sistema ou tentativas de acesso não autorizado. Se

os dados forem compartilhados entre vários usuários, o sistema deve evitar possíveis resultados anômalos devido à acessos concorrentes.

Explicação do modelo relacional.

Começar com um minimundo para dar contexto ao exemplo.

Minimundo: Sistema para gerenciamento de uma biblioteca.

- Usuários da biblioteca podem buscar livros por título, autor e assunto e o resultado da busca deve retornar o número de exemplares disponíveis.
- O sistema deve permitir o cadastro de livros, além da atualização das informações e remoção do livro do catálogo quando ele for danificado.
- Usuários podem alugar livros por 7 dias e o sistema deve registrar a data de início do empréstimo além da data de devolução.
- No momento da devolução, o sistema deve verificar se a entrega está atrasada e, se sim, exibir o valor total da multa (R\$ 1.50 por dia de atraso).
- A entrega só deve ser realizada mediante o pagamento da multa.
- Para cada usuário, o sistema deve registrar o cpf, nome, idade e telefone.

Perguntar aos alunos quais dados precisam ser armazenados pelo sistema.

Perguntar se eles fossem organizar esses dados usando excel e tentando ser bem organizados, como eles fariam.

Se ninguém lembrar dos aluguéis, mencionar. Se eles não lembrarem do usuário nos aluguéis, mencionar que precisamos saber quem alugou o livro para podermos cobrar se ele atrasar.

Desenhar a tabela livro e adicionar alguns itens.

Modelo de Dados Relacional [1]

Definições Básicas

- No modelo relacional, dados do mesmo tipo (livros, usuários, aluguéis) são armazenados em tabelas. Tabelas também são chamadas de relações (relations).
- Cada linha representa um “objeto” específico daquele tipo, i.e., uma **instância** da tabela. Linhas de tabelas são chamadas de **tuplas, registros (records)**.
- Tabelas possuem colunas que definem nomes únicos e os tipos dos valores armazenados. Colunas também são chamadas de atributos da tabela. *São parecidas com uma linha de cabeçalho em uma tabela do excel.*
- A coleção de informações armazenadas no banco de dados em um determinado momento é chamada de instância do banco de dados.
- O design geral do banco de dados é chamado de esquema do banco de dados. Os esquemas são alterados com pouca frequência, se é que são alterados. A relation (table) schema consists of a list of attributes and their corresponding domains.

Linguagens de Bancos de Dados.

São utilizadas para manipular o esquema, restrições, permissões e os dados do banco de dados.

SQL.

- IBM developed the original version of SQL, originally called Sequel, as part of the System R project in the early 1970s.
- SQL significa Structured Query Language.
- Many products support the SQL language. SQL has clearly established itself as the standard relational database language.

Criação de tabelas:

```
CREATE TABLE produto(id INTEGER, nome TEXT, preco REAL, estoque INTEGER);
```

Inserção de dados em tabelas:

```
INSERT INTO produto VALUES (3, 'Pasta de Dente', 4.30, 25);
```

```
INSERT INTO produto VALUES (5, 'Pizza', 16.50, 3);
```

```
INSERT INTO produto VALUES (4, 'Sabonete', 2.75, 123);
```

Fazer uma figura mostrando como a tabela ficaria depois destas inserções.

[Exercício] Solicitar que os alunos façam um exemplo de tabela Usuário com as informações do minimundo e adicionem alguns dados nesta tabela.

Linguagens de Bancos de Dados

Data-manipulation language (DML) is a language that enables users to access or manipulate data as organized by the appropriate data model. The types of access are Retrieval, Insertion, Deletion, and Modification.

- A query is a statement requesting the retrieval of information.

Data-definition language (DDL) is a language used to specify a database schema and database consistency constraints. Database systems implement integrity constraints that can be tested with minimal overhead:

- Domain Constraints. A domain of possible values must be associated with every attribute (for example, integer types, character types, date/time types).
- Referential Integrity. There are cases where we wish to ensure that a value that appears in one relation for a given set of attributes also appears in a certain set of attributes in another relation (referential integrity). For example, the department listed for each course must be one that actually exists.
- Assertions. An assertion is any condition that the database must always satisfy. Domain constraints and referential-integrity constraints are special forms of assertions. However, there are many constraints that we cannot express by using only these special forms. For example, "Every department must have at least five courses offered every semester" must be expressed as an assertion.
- Authorization. We may want to differentiate among the users as far as the type of access they are permitted on various data values in the database. These

differentiations are expressed in terms of authorization, the most common being: read authorization, which allows reading, but not modification, of data; insert authorization, which allows insertion of new data, but not modification of existing data; update authorization, which allows modification, but not deletion, of data; and delete authorization, which allows deletion of data.

Como criar um programa em Python capaz de executar comandos SQL

- Um script para criação do banco. Tipicamente é feito 1x só.
- *[exercício] Solicitar que os alunos criem o script para criar a tabela usuário. Deixar claro que a criação das duas tabelas podem e devem estar no mesmo script.*
- Outros scripts para manipulação de informações no banco. Operações de inserção, atualização, busca e remoção são feitas com frequência.

Na aula, faremos quatro scripts:

- Um que cria o banco de dados.
- Um para listar as tabelas e verificar se o banco foi criado corretamente.
- Outro para inserir informações.
- Outro para listar as informações inseridas para verificar se elas estão corretas.

Note que após a execução do primeiro script será criado um arquivo que irá armazenar o banco de dados.

[1] Silberschatz, A., Korth, H. F., & Sudarshan, S. (2002). Database system concepts (Vol. 5). New York: McGraw-Hill.

[2] Elmasri, R., Navathe, S. B., & Pinheiro, M. G. (2005). Sistemas de banco de dados.