



Bases de Dados

Aula 2: Modelo Relacional – Visão Geral e Terminologia

Prof. Eduardo Corrêa Gonçalves

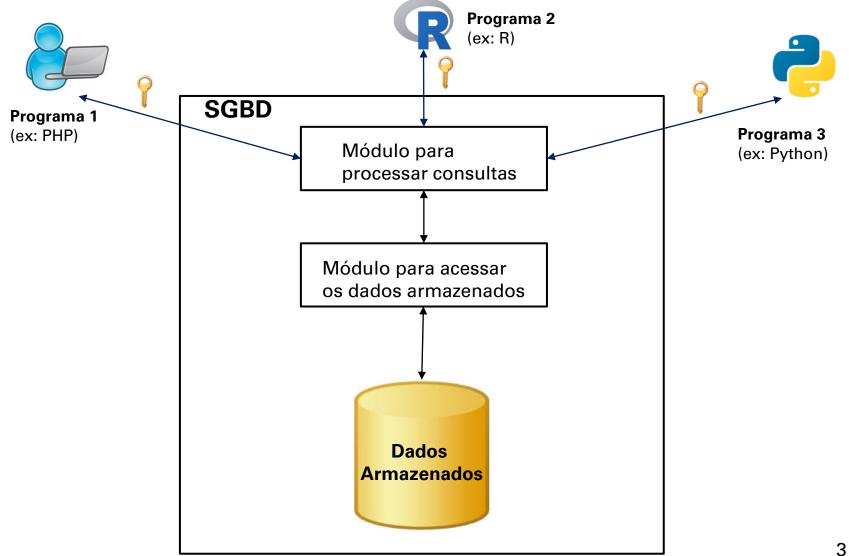


Tópicos

- Modelo Relacional
 - O que é um Modelo de Dados?
 - Introdução ao Modelo Relacional
 - Visão Geral
 - Terminologia
 - Definição de Esquemas



SGBDs (funcionamento simplificado)





SGBDs (funcionamento simplificado)

- SGBD: software responsável pelo gerenciamento completo de BDs.
 - Qualquer programa que queira acessar um BD (ou seja, o repositório de informações propriamente dito), tem que fazê-lo através do SGBD.
 - Apenas o SGBD possui acesso direto aos dados.
 - Os programas precisam de uma chave de acesso (ex: login, senha e endereço do servidor que hospeda o SGBD) para que possam realizar a conexão com um banco de dados gerenciado pelo SGBD.
 - Na verdade eles se conectam com o SGBD. O SGBD sabe quais os BDs ou partes de um BD aos quais o programa tem direito de acesso.
 - Um SGBD pode gerenciar vários BDs (ex: BD do sistema acadêmico da ENCE + BDs de cada pesquisa do IBGE).
 - Além de garantir a segurança dos dados, o SGBD é responsável por garantir o acesso eficiente e concorrente aos dados.



Modelo Relacional

- Modelo de dados: Notação adotada pelo SGBD para descrever dados ou informações.
 - Normalmente consiste em 3 partes:
 - Estrutura dos dados: estruturas utilizadas para realizar a representação conceitual dos dados (ex.: tabela, documento JSON, grafo, séries temporal, ...).
 - Operações sobre os dados: conjunto possível de operações de consulta (operações para recuperação de informações) e modificação (operações que modificam os dados) que podem ser executadas sobre o BD.
 - Restrições sobre os dados: forma pela qual indica-se as limitações sobre como os dados podem se apresentar. Podem variar desde algo muito simples ("um filme deve ter apenas um título") a restrições bem complexas.
 - Exemplos de modelos de dados adotados por SGBD's: Relacional,
 Orientado a Documentos (document store), Orientado a Grafos (graph dbms), Série Temporal (time series), ...



Modelo Relacional

- O **Modelo Relacional** é o mais antigo e ainda o mais largamente utilizado pelos SGBD's comerciais.
- Neste modelo:
 - Estrutura dos dados: baseado em tabelas.
 - Operações sobre os dados: definidas pela linguagem SQL.
 - Restrições sobre os dados: chave, integridade referencial, NOT NULL e outras.



Modelo Relacional

- Proposto nos anos 1970, o Modelo Relacional obteve grande sucesso principalmente devido aos seguintes motivos:
 - Possui uma base matemática sólida para representação e operações de consulta aos dados.
 - Essas operações são definidas pela Álgebra Relacional.
 - Foi o primeiro a propor a separação do armazenamento físico dos dados de sua representação conceitual.
 - Isso significa que: o usuário não precisa saber como os dados estão gravados de fato no disco (Árvore B, Lista Duplamente Encadeada, etc.); Ele sempre enxerga o BD como um simples conjunto de tabelas.
 - Introduziu uma linguagem de consulta de alto nível para a manipulação dos dados: a SQL.
 - Antes do modelo relacional, os usuários precisavam criar programas para manipular os dados.



RELAÇÃO

- O modelo relacional nos fornece uma forma única para representar os dados: uma tabela bidimensional denominada relação.
- Ex: Filme

Γ	titulo	ano	resumo	pais	duracao	avaliacao
1	Monty Python em Busca do Cálice Sagrado	1975	King Arthur and his knights embark on a low-budget search for the Grail	UK	91	8.30000
2	Cidade de Deus	2002	Two boys growing up in a violent neighborhood of Rio de Janeiro take di	BR	130	8.69999
3	Um Sonho de Liberdade	1994	Two imprisoned men bond over a number of years, finding solace and e	US	142	9.30000
4	Intocáveis	2011	After he becomes a quadriplegic from a paragliding accident, an aristocr	FR	112	8.6
	As Pontes de Madison	1995	The path of Francesca Johnson seems destined when an unexpected for	US	135	7.3

- Cada linha da relação representa um único filme.
- As colunas representam propriedade dos filmes.
- Nos SGBDs, as relações também são chamadas de tabelas.



ATRIBUTOS

- As colunas de uma relação são denominadas atributos.
- **Ex**.: Na relação *Filme*, os atributos são:

```
"título";
"ano";
"resumo";
"país";
"duração";
"avaliação".
```

 Nos SGBDs, um atributo também é chamado de campo (field) ou simplesmente de coluna.

Γ	titulo	ano	resumo	pais	duracao	avaliacao
1	Monty Python em Busca do Cálice Sagrado	1975	King Arthur and his knights embark on a low-budget search for the Grail	UK	91	8.30000
2	Cidade de Deus	2002	Two boys growing up in a violent neighborhood of Rio de Janeiro take di	BR	130	8.69999
3	Um Sonho de Liberdade	1994	Two imprisoned men bond over a number of years, finding solace and e	US	142	9.30000
4	Intocáveis	2011	After he becomes a quadriplegic from a paragliding accident, an aristocr	FR	112	8.6
5	As Pontes de Madison	1995	The path of Francesca Johnson seems destined when an unexpected for	US	135	7.3



TUPLAS

- A linhas de uma relação (excetuando aquelas contendo os nomes dos atributos) são chamadas de **tuplas**.
- Uma tupla contém um componente (valor) para cada atributo da relação.
- Notação: conjunto com n-valores (na ordem default), um para cada atributo.
 - $t = (V_1, V_2, ..., V_n)$
- Ex.: representação da quarta tupla da relação Filme
 - *t* = ('Intocáveis', 2011, 'After he becomes...', 'FR',112, 8.6)

Γ	titulo	ano	resumo	pais	duracao	avaliacao
1	Monty Python em Busca do Cálice Sagrado	1975	King Arthur and his knights embark on a low-budget search for the Grail	UK	91	8.30000
2	Cidade de Deus	2002	Two boys growing up in a violent neighborhood of Rio de Janeiro take di	BR	130	8.69999
3	Um Sonho de Liberdade	1994	Two imprisoned men bond over a number of years, finding solace and e	US	142	9.30000
4	Intocáveis	2011	After he becomes a quadriplegic from a paragliding accident, an aristocr	FR	112	8.6
5	As Pontes de Madison	1995	The path of Francesca Johnson seems destined when an unexpected for	US	135	7.3



TUPLAS

 Nos SGBDs, uma tupla também é chamada de registro (record) ou simplesmente de linha.

	titulo	ano	resumo	pais	duracao	avaliacao
1	Monty Python em Busca do Cálice Sagrado	1975	King Arthur and his knights embark on a low-budget search for the Grail	UK	91	8.30000
2	Cidade de Deus	2002	Two boys growing up in a violent neighborhood of Rio de Janeiro take di	BR	130	8.69999
3	Um Sonho de Liberdade	1994	Two imprisoned men bond over a number of years, finding solace and e	US	142	9.30000
4	Intocáveis	2011	After he becomes a quadriplegic from a paragliding accident, an aristocr	FR	112	8.6
5	As Pontes de Madison	1995	The path of Francesca Johnson seems destined when an unexpected for	US	135	7.3



ESQUEMAS

- O nome da relação e o seu conjunto de atributos é denominado esquema (schema) da relação.
- Notação: nome da relação seguido por seus atributos entre parênteses.
 - $R(A_1, A_2, ..., A_n)$
- Exemplo:
 - Filme (titulo, ano, resumo, pais, duracao, avaliacao)

Γ	titulo	ano	resumo	pais	duracao	avaliacao
l	Monty Python em Busca do Cálice Sagrado	1975	King Arthur and his knights embark on a low-budget search for the Grail	UK	91	8.30000
12	Cidade de Deus	2002	Two boys growing up in a violent neighborhood of Rio de Janeiro take di	BR	130	8.69999
1	Um Sonho de Liberdade	1994	Two imprisoned men bond over a number of years, finding solace and e	US	142	9.30000
4	Intocáveis	2011	After he becomes a quadriplegic from a paragliding accident, an aristocr	FR	112	8.6
	As Pontes de Madison	1995	The path of Francesca Johnson seems destined when an unexpected for	US	135	7.3



ESQUEMAS

- Os atributos da relação representam um conjunto e não uma lista.
- Entretanto, normalmente especificamos uma ordem padrão para estes atributos, pois assim torna-se mais fácil trabalhar na prática.
- O nome da relação é especificado com a primeira letra em maiúsculo. E os nomes dos atributos em minúsculo.
 - Filme(titulo, ano, resumo, pais, duracao, avaliacao)
- No modelo relacional, um banco de dados é formado por uma ou mais relações. O conjunto de esquemas do banco é chamado de esquema relacional do banco ou simplesmente esquema do banco.

Γ	titulo	ano	resumo	pais	duracao	avaliacao
1	Monty Python em Busca do Cálice Sagrado	1975	King Arthur and his knights embark on a low-budget search for the Grail	UK	91	8.30000
2	Cidade de Deus	2002	Two boys growing up in a violent neighborhood of Rio de Janeiro take di	BR	130	8.69999
3	Um Sonho de Liberdade	1994	Two imprisoned men bond over a number of years, finding solace and e	US	142	9.30000
4	Intocáveis	2011	After he becomes a quadriplegic from a paragliding accident, an aristocr	FR	112	8.6
5	As Pontes de Madison	1995	The path of Francesca Johnson seems destined when an unexpected for	US	135	7.3



Modelo Relacional: Terminologia

INSTÂNCIA

- Chamamos o conjunto de tuplas de uma dada relação de instância da relação.
- IMPORTANTE: Relações são conjuntos de tuplas e não listas de tuplas.
 Sendo assim, a ordem em que as tuplas de uma relação são apresentadas é "imaterial".
 - Basicamente, qualquer ordem de apresentação das tuplas da relação Filme não vai modificar o conteúdo da relação Filme.

Γ	titulo	ano	resumo	pais	duracao	avaliacao
ı	1 Monty Python em Busca do Cálice Sagrado	1975	King Arthur and his knights embark on a low-budget search for the Grail	UK	91	8.30000
ı	2 Cidade de Deus	2002	Two boys growing up in a violent neighborhood of Rio de Janeiro take di	BR	130	8.69999
l	3 Um Sonho de Liberdade	1994	Two imprisoned men bond over a number of years, finding solace and e	US	142	9.30000
ŀ	4 Intocáveis	2011	After he becomes a quadriplegic from a paragliding accident, an aristocr	FR	112	8.6
ı	5 As Pontes de Madison	1995	The path of Francesca Johnson seems destined when an unexpected for	US	135	7.3



Modelo Relacional: Terminologia

INSTÂNCIA

- Uma relação sobre filmes não é estática; Por exemplo: a cada dia filmes novos são inseridos no IMDb.
 - Desta forma, espera-se que a instância da relação mude com bastante frequência.
- De maneira oposta, o esquema da relação costuma ser imutável ou muda com pouca frequência.
 - Apenas de tempos em tempos pode ser decidido acrescentar um atributo, ou renomear a relação ou um atributo, mudar um tipo de atributo, etc.



TIPOS DE DADOS

- Os atributos de uma relação podem ser de diferentes tipos, tais como alfanumérico, inteiro e real, entre outros.
- Na relação Filme, existem:
 - dois atributos inteiros ("ano" e "duracao");
 - um atributo real ("avaliacao");
 - três atributos alfanuméricos ("titulo", "resumo" e "pais").

Γ	titulo	ano	resumo	pais	duracao	avaliacao
1	Monty Python em Busca do Cálice Sagrado	1975	King Arthur and his knights embark on a low-budget search for the Grail	UK	91	8.30000
2	2 Cidade de Deus	2002	Two boys growing up in a violent neighborhood of Rio de Janeiro take di	BR	130	8.69999
3	Um Sonho de Liberdade	1994	Two imprisoned men bond over a number of years, finding solace and e	US	142	9.30000
4	Intocáveis	2011	After he becomes a quadriplegic from a paragliding accident, an aristocr	FR	112	8.6
	As Pontes de Madison	1995	The path of Francesca Johnson seems destined when an unexpected for	US	135	7.3



TIPOS DE DADOS

- É possível incluir o tipo de dado para cada atributo ao fazer a representação de um esquema.
 - Basta utilizar o símbolo ": " (dois-pontos) para separar o nome do atributo e o tipo.

Exemplo:

Filme (titulo:alfanumérico, ano:inteiro, resumo:alfanumérico, pais:alfanumérico, duracao:inteiro, avaliacao:real)

Γ	titulo	ano	resumo	pais	duracao	avaliacao
ŀ	Monty Python em Busca do Cálice Sagrado	1975	King Arthur and his knights embark on a low-budget search for the Grail	UK	91	8.30000
2	Cidade de Deus	2002	Two boys growing up in a violent neighborhood of Rio de Janeiro take di	BR	130	8.69999
:	Um Sonho de Liberdade	1994	Two imprisoned men bond over a number of years, finding solace and e	US	142	9.30000
4	Intocáveis	2011	After he becomes a quadriplegic from a paragliding accident, an aristocr	FR	112	8.6
!	As Pontes de Madison	1995	The path of Francesca Johnson seems destined when an unexpected for	US	135	7.3



DOMÍNIOS

- Também é assumido que associado a cada atributo de uma relação exista um domínio.
- Por exemplo: para o atributo "avaliacao", que armazena a média das avaliações atribuídas pelo público para cada filme, o domínio é representado pelos números reais compreendidos entre 0,0 e 10,0.
 - Valores como 10,5 ou -1 representam violações ao domínio de "avaliacao", não podendo ser aceitos pelo SGBD.
 - Mas, para implementar a checagem do domínio, é preciso associar uma restrição ao campo, conforme veremos na Aula 05.

Γ	titulo	ano	resumo	pais	duracao	avaliacao
ŀ	Monty Python em Busca do Cálice Sagrado	1975	King Arthur and his knights embark on a low-budget search for the Grail	UK	91	8.30000
1	2 Cidade de Deus	2002	Two boys growing up in a violent neighborhood of Rio de Janeiro take di	BR	130	8.69999
:	Um Sonho de Liberdade	1994	Two imprisoned men bond over a number of years, finding solace and e	US	142	9.30000
1	1 Intocáveis	2011	After he becomes a quadriplegic from a paragliding accident, an aristocr	FR	112	8.6
1	As Pontes de Madison	1995	The path of Francesca Johnson seems destined when an unexpected for	US	135	7.3



DOMÍNIOS

- Os valores associados ao domínio de um atributo devem ser sempre atômicos (simples, indivisíveis).
- Portanto, um componente qualquer em uma tupla deve representar um único valor do domínio do atributo
 - Não pode armazenar um vetor ou uma lista de valores, por exemplo.

Γ	titulo	ano	resumo	pais	duracao	avaliacao
ŀ	Monty Python em Busca do Cálice Sagrado	1975	King Arthur and his knights embark on a low-budget search for the Grail	UK	91	8.30000
1	2 Cidade de Deus	2002	Two boys growing up in a violent neighborhood of Rio de Janeiro take di	BR	130	8.69999
:	Um Sonho de Liberdade	1994	Two imprisoned men bond over a number of years, finding solace and e	US	142	9.30000
1	1 Intocáveis	2011	After he becomes a quadriplegic from a paragliding accident, an aristocr	FR	112	8.6
1	As Pontes de Madison	1995	The path of Francesca Johnson seems destined when an unexpected for	US	135	7.3



CHAVE (Key)

- Existem diferentes tipos de restrições que o modelo relacional nos permite atribuir a esquemas. Este assunto será abordado na Aula 05.
- Entretanto, há um tipo de restrição que é fundamental e que iremos apresentar agora: restrição de chave (key constraint).
- Um conjunto de atributos forma a chave para uma relação se nós não permitimos que duas tuplas da instância da relação armazenem o mesmo valor em todos os atributos que formam a chave.



CHAVE

- Exemplo: na relação Filme, podemos declarar uma chave formada por dois atributos: titulo e ano.
 - Isto porque, como projetistas do BD, não acreditamos que existam dois filmes lançados com o mesmo título no mesmo ano.
 - Observe que apenas o título não poderia ser usado para formar uma chave, pois muitas vezes ocorrem refilmagens (ex.: Caça-Fantasmas original de 1984 e Caça-Fantasmas de 2016).
 - Obviamente, apenas o ano não pode formar uma chave, pois existem muitos filmes produzidos em um mesmo ano.



CHAVE

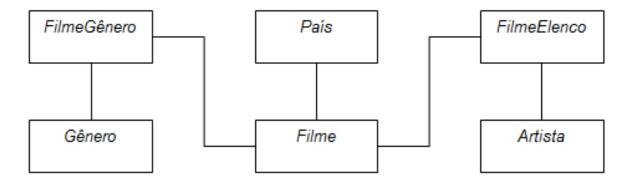
- Indicamos os atributos que formam a chave de uma relação sublinhando os mesmos na definição do esquema.
 - Filme (titulo, ano, resumo, pais, duracao, avaliacao)

	titulo	ano	resumo	pais	duracao	avaliacao
1	Monty Python em Busca do Cálice Sagrado	1975	King Arthur and his knights embark on a low-budget search for the Grail	UK	91	8.30000
2	Cidade de Deus	2002	Two boys growing up in a violent neighborhood of Rio de Janeiro take di	BR	130	8.69999
3	Um Sonho de Liberdade	1994	Two imprisoned men bond over a number of years, finding solace and e	US	142	9.30000
4	Intocáveis	2011	After he becomes a quadriplegic from a paragliding accident, an aristocr	FR	112	8.6
5	As Pontes de Madison	1995	The path of Francesca Johnson seems destined when an unexpected for	US	135	7.3



Exemplo: Esquema do BD de Filmes

- Em um BD relacional tipicamente existirão **muitas relações**. Algumas delas estarão **associadas** (ligadas, vinculadas)
 - Os motivos para tal serão devidamente esclarecidos em aulas posteriores. Por enquanto apenas considere que este é um fato.
 - A figura abaixo apresenta um exemplo de projeto de BD que poderia ser utilizado em um web site sobre cinema.

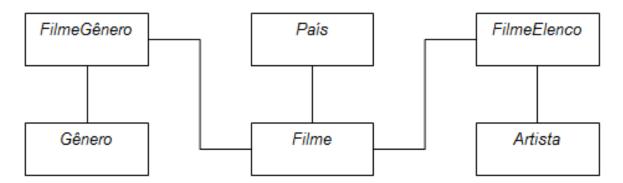


- Neste esquema de banco, cada relação é representada por uma caixa (devidamente identificada) e cada associação entre pares de relações é representada por uma linha unindo as duas caixas.
- Dessa maneira, o projeto é composto por seis relações e cinco associações.



Exemplo: Esquema do BD de Filmes

- Filme: informações sobre os filmes. Chave = titulo e ano
- Pais: países em que filme pode ser produzido. Chave = sigla do país.
- Artista: informações sobre os artistas que atuam nos filmes. Chave = nome (ok, não é uma boa chave, mas por questão de simplicidade, vamos utilizá-la inicialmente).
- Genero: conjunto de possíveis gêneros para um filme. Chave = nome.
- *FilmeGenero*: gêneros associados aos filmes da relação *Filme*. Chave = titulo do filme, ano e nome do gênero.
- FilmeElenco: artistas que atuaram nos filmes da relação Filme. Chave = título do filme, ano e nome do artista.





Exemplo: Esquema do BD de Filmes

- Representação das relações:
 - Filme (titulo, ano, resumo, pais, duracao, avaliacao)
 - Pais (sigla, nome)
 - Artista (nome, sexo)
 - Genero (nome)
 - FilmeGenero (titulofilme, ano, nomegenero)
 - FilmeElenco (titulofilme, ano, nomeartista)



- Os bancos de dados relacionais são manipulados através de uma linguagem padrão, desenvolvida especialmente para o ambiente relacional, denominada SQL (Structured Query Language).
- As principais instruções da SQL são apresentadas na tabela abaixo:

Finalidade	Instruções	Descrição
Recuperação de Dados	SELECT	Recupera registros armazenados em tabelas do banco de dados.
Manipulação de Dados	INSERT DELETE UPDATE	Inserção, alteração e remoção de registros de tabelas do banco de dados. Este subconjunto de instruções da SQL é conhecido como DML (<i>Data Manipulation Language</i>).
Definição de Dados	CREATE ALTER DROP	Criação, alteração e exclusão de objetos do banco de dados (ex.: tabelas, índices, etc.). Este subconjunto de instruções é conhecido como DDL (<i>Data Definition Language</i>).
Controle de Transações	COMMIT ROLLBACK	Gerenciam as modificações realizadas pelos comandos DML. Permitem agrupar as alterações dos dados em transações lógicas.



- Nesta aula e na próxima, será apresentada a forma básica para declarar (criar) tabelas utilizando a instrução CREATE TABLE.
 - Trata-se, obviamente, de uma instrução DDL!
 - Em sua forma básica, a declaração de uma tabela consiste:
 - No uso da palavra CREATE TABLE seguida pelo nome da relação; e
 - Uma lista de atributos e seus respectivos tipos. Esta especificação deve ser feita entre parênteses, com cada definição de atributo separada por vírgula;
 - Os atributos que compõem a chave são indicados com o uso da palavra PRIMARY KEY (chave primária).

```
CREATE TABLE Pais (
sigla CHAR(2),
nome VARCHAR(50),
PRIMARY KEY (sigla)
);
```



- O exemplo ao lado ilustra o comando CREATE TABLE utilizado para criar a tabela *Filme* no SQLite.
 - Os atributos titulo e resumo são do tipo VARCHAR(n) (alfanumérico com tamanho máximo n).
 - O atributo pais é do tipo CHAR(n) (alfanumérico com tamanho fixo n).
 - Os atributos ano e duracao do tipo INT (número inteiro no SQLite).
 - Por fim, avaliacao é do tipo NUM (número real no SQLite).
 - A chave da relação é composta pelos atributos titulo e ano.

```
CREATE TABLE Filme (
titulo VARCHAR(100),
ano INT,
resumo VARCHAR(2000),
pais CHAR(2),
duracao INT,
avaliacao NUM,
PRIMARY KEY (titulo, ano)
):
```



Script completo para criar todas as relações apresentadas no Slide 25.

```
CREATE TABLE Filme (
titulo VARCHAR(100).
ano INT,
resumo VARCHAR(2000),
pais CHAR(2),
duração INT,
avaliação NUM,
PRIMARY KEY (titulo, ano)
);
CREATE TABLE Pais (
sigla CHAR(2),
nome VARCHAR(60),
PRIMARY KEY (sigla)
);
CREATE TABLE Artista (
nome VARCHAR(60),
sexo CHAR(1),
PRIMARY KEY (nome)
);
```

```
CREATE TABLE Genero (
nome VARCHAR(15),
PRIMARY KEY (nome)
);
CREATE TABLE FilmeGenero (
titulofilme VARCHAR(100),
ano INT,
nomegenero VARCHAR(100),
PRIMARY KEY (titulofilme,
ano, nomegenero)
);
CREATE TABLE FilmeElenco (
titulofilme VARCHAR(100),
ano INT,
nomeartista VARCHAR(60),
PRIMARY KEY (titulofilme,
ano, nomeartista)
);
```



Próximas Aulas

Na Aula 03

Praticaremos a criação de esquemas com o uso da SQL.

Na Aula 04

 Aprenderemos a inserir tuplas, excluir tuplas e atualizar tuplas de relações criadas, utilizando a SQL.

Na Aula 05

- Abordaremos os diferentes tipos restrições e aprenderemos a criar tabelas contendo chaves estrangeiras.
- Começaremos a apresentar o motivo de os SGBDs relacionais utilizarem muitas tabelas para guardarem os dados (em vez de uma só).