

## **4-Compreender e projetar banco de dados (DL)**

**E-Book - Apostila**

# Compreender e projetar banco de dados

Que mundo fascinante este no qual vivemos!

Em 2017, a "The Economist", em reportagem de capa, dizia Data is the new oil, ou seja, **Dados são o novo petróleo**. Estamos falando que é um bem tão valioso que precisa ser explorado, cuidado e tratado de maneira correta para que entregue o seu valor.



**Saiba Mais**

## DADO É O NOVO PETRÓLEO

Para embasar o que estamos falando, sugerimos você assistir ao vídeo a seguir para saber o que é esse novo petróleo.

<https://www.youtube.com/watch?v=t4Bbp1PWZaE>

O ser humano desde os primórdios da sua existência tem necessidade de deixar registrado como viveu na Terra. E as formas de registro são variadas. Na figura vemos as pinturas pré-históricas em cavernas (chamadas de arte rupestre) que ilustram o que os historiadores dizem ser cenas de caça, rituais, o cotidiano, expressando uma espécie de linguagem visual, conceitos, símbolos, valores e crenças - ou seja, o registro do povo da época:

A Figura anexa mostra uma pedra rochosa com alguns desenhos, que podem ser interpretados como corpos humanos e de animais.

Atualmente, há avanços tecnológicos em todos os aspectos no nosso dia a dia: telefones celulares, televisão via satélite, aeronaves avançadas, medicina moderna e computadores. Negócios de todos os tipos e tamanhos contam com computadores e os sistemas de informações dão suporte em um grau que teria sido inimaginável há apenas alguns anos.

Estamos falando que o ser humano da sociedade moderna está deixando de armazenar seus registros em meio físico, e cada vez mais digitaliza sua vida: fotos de viagem, conversas com amigos, documentos, tudo atualmente é transferido do mundo real para o virtual.

As empresas cada vez mais desenvolvem meios para esse novo mundo de possibilidades que é o universo digital e promovem cada vez mais o uso do meio digital.

Sabemos que a informação é um importante capital da empresa, pois o conhecimento que as corporações têm de seus clientes é muito valioso e é considerado o novo petróleo do mundo.

Por isso, para diminuir a redundância e o tempo, as organizações optam por salvar cada vez mais dados dos seus usuários de diferentes formas e utilizar este conhecimento de maneira oportuna.

O planeta gera informação a todo momento: a cada 60 segundos são gerados um grande volume de dados nas redes sociais, conforme mostra a figura a seguir, os dados são correspondentes a junho de 2020:

Devido a este grande volume de informação, foi assinada a Lei n. 13.709/2018 para proteger as informações pessoais de consumidores e usuários de serviços, conhecida como **Lei Geral de Proteção de Dados (LGPD)**, a qual prevê que o usuário que sofrer vazamento dos seus dados possa ser indenizado e organizações poderão ser multadas por falha de segurança.

Devido ao volume de dados produzidos pela humanidade, vemos a importância e a relevância desses dados para as empresas continuarem sendo competitivas.

Consequentemente, temos a necessidade da proteção destes dados que são produzidos, por isso, estudaremos sobre como guardar essas informações, para que seja gerado um valor para a empresa. Nesta unidade, buscaremos responder aos seguintes questionamentos:

- Como encontrar os dados que atendam a necessidade da empresa?
- Como documentar de forma eficiente os requisitos?
- Como pensar em dados de forma organizada?

- Como projetar os dados para que sejam armazenados de maneira eficiente?
- Como podemos recuperar estes dados com eficácia?



### Ao final deste conteúdo, você será capaz de:

- Identificar as diferenças entre os conceitos de dado, informação e conhecimento.
- Relacionar os componentes e objetivos de um sistema de banco de dados.
- Explicar como os dados são modelados ao longo do processo de desenvolvimento de software.



Bons estudos!



# Dados x

# Informação

Até aqui, falamos sobre dados e informação. Mas você sabe a diferença?

Há muita dúvida sobre o que é dado e o que é informação, pois ambos os conceitos estão relacionados. **Dados** são fatos brutos, que ainda não foram processados para revelar seu significado e devem ser formatados adequadamente para o armazenamento, o processamento e a apresentação.

**Dados** são fatos que podem ser gravados e que possuem um significado implícito (ELMASRI, 2011).

Já **Informação** é o resultado do processamento, manipulação e organização dos dados de tal forma que possa representar uma modificação (quantitativa ou qualitativa) no conhecimento do sistema (pessoa, animal, objeto ou máquina) que a recebe. Informação é o conjunto de dados associados a um contexto, de maneira que seja possível interpretá-la e analisá-la para produzir conhecimento e/ou tomar decisões (PUGA; FRANÇA; GOYA, 2013).

O que é feito com a informação tratada gera **CONHECIMENTO**.

**Conhecimento** é quando temos familiaridade, consciência e compreensão destas informações e através deste conhecimento é possível gerar algo novo.

Um exemplo é: Dado x Informação e Conhecimento apresentado na tabela a seguir:

**TABELA 1 - Dado, Informação e Conhecimento**

DADO	INFORMAÇÃO	CONHECIMENTO
NOME	Marizildo	Funcionário da empresa XPTO

INSTITUCIONAL, 2022.

# Classificação de dados

Os dados podem ser classificados de duas maneiras: pela **fonte de dados** ou pelo **formato de dados**.

## Fonte de dados

Quanto à fonte dos dados, que trata do local onde os dados são encontrados, podemos classificar como dados internos ou dados externos.

Os **Dados Internos** são os dados estruturados, previsíveis, permanentes e fáceis de obter. Exemplo: Sistemas ERP, CRM, BI, SCM, RH.

Os **Dados Externos** são os dados não estruturados, randômicos, voláteis e difíceis de obter. Exemplo: mídia social, automação, Saas, Paas, APIs, Análise de Textos, Reconhecimento de Voz, Processamento de Linguagem Natural e Internet das Coisas.

## Formato de dados

Quanto ao formato dos dados, que trata da apresentação, como os dados surgem, podemos classificar como dados estruturados, dados semiestruturados ou dados não estruturados.

### FIGURA 1 - Formato de dados



INSTITUCIONAL, 2022.

A Figura representa graficamente os três formatos de dados: estruturados, semiestruturados e não estruturados.

Os **Dados Estruturados** são organizados em linhas e colunas, geralmente presentes nos bancos de dados relacionais, são dados de fácil processamento e recuperação. Esses dados possuem limitações quando precisam tratar grandes volumes de dados. Exemplo: bancos de dados relacionais, Datawarehouse.

Os **Dados Semiestruturados** são dados que apesar de ter propriedades internas organizadas, não podem ser armazenados em um banco de dados.

undefined

Os **Dados Não Estruturados** não podem ser organizados em linhas e colunas, como vídeos, comentários em redes sociais e e-mails, dentre outros. Geralmente são dados de difícil acesso e recuperação e muitas vezes não dispõem de componentes necessários para identificação de tipo de processamento. Esse tipo de dado foi criado por causa do Big Data (veremos mais adiante).

# Banco de Dados

O que você sabe sobre banco de dados?

**Banco de Dados** é um conjunto de dados integrados que tem por objetivo atender a uma comunidade de usuários (HEUSER, 2011).

Acompanhe algumas definições para banco de dados e que são complementares:

Banco de dados é uma coleção de dados armazenados e organizados de modo a atender as necessidades integradas dos seus usuários. Possibilita a consulta e a manipulação dos dados, podendo ser manual ou computadorizado (PUGA; FRANÇA; GOYA, 2013).

A definição de banco de dados também se amplia ao entender que é uma coleção de dados relacionados. Representa algum aspecto do mundo real, às vezes chamado de minimundo ou universo de discurso (UoD - Universe of Discourse). As mudanças no minimundo são refletidas no banco de dados (ELMASRI; NAVATHE, 2018).

Banco de dados é projetado, construído e populado com dados para uma finalidade específica. Ele possui um grupo definido de usuários e algumas aplicações previamente concebidas nas quais esses usuários estão interessados (ELMASRI; NAVATHE, 2018).

Um banco de dados pode ter qualquer tamanho e complexidade e não necessariamente precisa de um dispositivo eletrônico (computador) para existir.

**Exemplo:** uma lista de nomes e endereços. Na figura 4 há alguns exemplos do uso do banco de dados em um Controle Bancário, Controle de Livros ou ainda em um Controle de Locação.

## FIGURA 2 - Exemplos de Bancos de Dados



Esses bancos de dados são organizados com a estrutura básica de tabelas, que representam algo do mundo real em que há necessidade de conhecer e armazenar essas informações e colunas, que representam uma característica, propriedade dessa tabela. Cada conjunto de colunas preenchidas será uma linha da tabela, um registro armazenado no banco de dados, conforme exemplifica a tabela a seguir:

## TABELA 2 - Exemplo Tabela, Atributo e Registro

FUNCIONÁRIO	
Nome	Salário
Cariolando dos Anjos	R\$ 1.045,00

INSTITUCIONAL, 2022.

## Funcionalidades do Banco de Dados

Os bancos de dados possuem um conjunto de funcionalidades:

### Controle de sincronização

Pela inclusão de mecanismos eficientes e eficazes que permitam o acesso ao banco de dados multiusuário tal que não haja interferência entre as distintas transações concomitantes que têm acesso aos dados compartilhados;

### Proteção e recuperação de falhas

Pela inclusão de mecanismos que protejam a base de dados fisicamente, evitando danos custosos em caso de pane e provendo procedimentos de recuperação automática das bases quando isto ocorrer;

### Segurança e acesso

Pela inclusão de mecanismos que permitam acesso selecionado a usuários autorizados e que mantenham controle e registro de operações críticas;

### Integridade dos dados

Pela inclusão de mecanismos que controlem e mantenham as restrições de integridade lógica dos dados, mantendo-os íntegros e corretos;

### Processamento de consultas

Pela inclusão de mecanismos que efetuem o processamento de consultas ao banco de dados através de linguagens de uso geral.

# Redundância de dados

A **redundância de dados** ocorre quando os dados são armazenados de forma desnecessária em locais diferentes, o que prejudica a manipulação dos dados, pois cada vez que o dado for alterado haverá retrabalho, localizando todos os lugares em que o mesmo se encontra. A Figura 5 esquematiza a redundância de dados, onde um sistema de cobrança, um sistema de cadastro e um sistema de frequência escolar possuem cada um o seu banco de dados contendo informações sobre estudantes:

**FIGURA 3 - Exemplo de Redundância de Dados**



Este tipo de abordagem também é conhecido como **processamento de arquivos**, no qual cada usuário define e implementa os arquivos necessários para uma aplicação de software específica como parte da programação da aplicação (ELMASRI; NAVATHE, 2018). E para resolver o problema de redundância de dados, é feito o **compartilhamento dos dados**, onde as informações são armazenadas em um único local e acessada por vários sistemas, demonstrado na Figura a seguir no qual um sistema de cobrança, um sistema de cadastro e um sistema de frequência escolar todos acessam o mesmo banco de dados contendo informações sobre estudantes:

**FIGURA 4 - Exemplo de Compartilhamento de Arquivos**



INSTITUCIONAL, 2022.

As principais características de um banco de dados com compartilhamento de dados são (ELMASRI; NAVATHE, 2018):

- Autodescrição de um sistema de banco de dados;
- Isolamento entre programas e dados e abstração de dados;
- Suporte de múltiplas visões dos dados;
- Compartilhamento de dados e processamento de transação multiusuário.

## Independência dos dados

O dado é o elemento mais estável de um sistema de informação.

No processamento de arquivos tradicional, a estrutura dos arquivos de dados será embutida nos programas de aplicação, de modo que quaisquer mudanças na estrutura de um arquivo podem exigir mudar todos os programas que acessam esse arquivo. Caso contrário, os programas de acesso ao SGBD (Sistema Gerenciador de Banco de Dados) não exigem essas mudanças na maioria dos casos. A estrutura dos arquivos de dados é armazenada no catálogo do SGBD separadamente dos programas de acesso. A esta propriedade chamamos de independência de dados do programa. Abstração de dados é uma característica que permite a independência dos dados em relação aos programas. (ELMASRI; NAVATHE, 2018).

O valor, o conteúdo do dado, se altera, mas o dado dificilmente sofre mudanças. Porém, o processo que, por exemplo, calcula o valor do salário líquido do funcionário pode sofrer modificações devido a uma alteração da lei do governo, por exemplo.

Há interatividade entre o dado e o processo, mas são independentes. Independência de (aplicações em relação aos) dados significa que uma alteração no modelo de dados afeta “pouco” as aplicações.

Exemplo de sistema com pouca independência: Modificação no esquema exige recompilação de todos os programas. As aplicações não contêm descrições dos dados. Com isso, a base de dados pode ser alterada sem que as aplicações sejam modificadas.

# Restrições de integridade

Há três tipos de restrições de integridade: **integridade referencial, domínio e semântica**.

## Integridade referencial

Quando a chave estrangeira contém um valor, esse valor se refere a uma Tupla (linha) válida existente em outra relação. Assegura que um valor que aparece em uma relação (tabela) para um determinado conjunto de atributos apareça em outro conjunto de atributos em outra relação (tabela). Como as tabelas em essência são relações, utilizam-se os termos matemáticos relação e tupla, no lugar de tabela e linhas.

Exemplo: se o curso de Computação é um registro que aparece na tabela estudante então deve existir um registro chamado Computação na tabela curso.

A integridade referencial possui características que devem ser atendidas conforme a tabela a seguir:

**TABELA 3 - Características da Integridade Referencial**

INTEGRIDADE REFERENCIAL	DESCRIÇÃO
-------------------------	-----------

<b>EXIGÊNCIA</b>	Uma chave estrangeira pode ter uma entrada nula, desde que não faça parte da chave primária de uma tabela que esteja relacionada. Ou seja, todo valor não nulo de chave estrangeira DEVE referenciar um valor de chave primária EXISTENTE.
<b>FINALIDADE</b>	É possível que um atributo NÃO tenha um valor correspondente, mas é impossível que tenha uma entrada inválida. A aplicação da regra de integridade referencial torna impossível a exclusão de uma linha em uma tabela cuja chave primária tenha valores obrigatórios de chave estrangeira em outra tabela.

INSTITUCIONAL, 2022.

## Domínio

As restrições de integridade resguardam o Banco de Dados contra danos acidentais, assegurando que mudanças feitas por usuários autorizados não resultem na perda de consistência de dados. Restrições de domínio são a forma mais elementar de restrições de integridade. Estas testam valores inseridos no Banco de Dados, e testam (efetuam) consultas para assegurar que as comparações façam sentido.

# Semântica

Há muitas restrições de integridade que não se encaixam nas categorias básicas. Essas restrições são chamadas de restrições semânticas (ou regras de negócio). Também chamada de regras do negócio.

Para aprender mais, temos convênio com a **Oracle Academy** e há um curso chamado **DataBase Foundations** que tem o objetivo de descrever as funcionalidades de um banco de dados relacional, usar a modelagem de dados para construir um banco de dados, desenvolver um diagrama entidade-relacionamento para modelar dados, mapear um modelo físico, entre outros.

## Sistema gerenciados de Banco de Dados

No início, as linguagens de programação eram responsáveis por definir e manipular os dados. Ou seja, era necessário programar o acesso aos dados. Isso mudou com o surgimento do Sistemas Gerenciador do Banco de Dados (SGBD) ou DataBase Management System (DBMS).

O **Sistema Gerenciador de Banco de Dados** (SGBD) tem o objetivo de manter grandes repositórios compartilhados de dados, ou seja, manter bancos de dados. É um software que incorpora as funções de definição, recuperação e alteração de dados em um banco de dados (HEUSER, 2011). A figura a seguir esquematiza o papel do SGBD no sistema de computação, onde as aplicações acessam o SGBD e o SGBD acessa o banco de dados.

## FIGURA 5 - SGBD



INSTITUCIONAL, 2022.

## Vantagens no uso de SGBD

O uso do sistema gerenciador de banco de dados tem as seguintes vantagens:

- Aumento na produtividade do usuário final;
- Minimização da inconsistência dos dados;
- Melhoria na integridade dos dados;
- Melhoria na tomada de decisão;
- Melhoria no compartilhamento de dados;
- Melhoria no acesso aos dados.

# Características de um SGBD

Para que um software de banco de dados seja considerado um SGBD deve possuir TODAS as características a seguir, do contrário é somente um banco de dados. As características são:

- Abstração (o usuário não precisa saber os detalhes como os dados são manipulados ou armazenados no banco de dados);
- Controle das Transações (garante a integridade dos dados);
- Controle de Concorrência (não permite que dois bancos de dados acessem os dados ao mesmo tempo);
- Independência física (dados não dependem do esquema, dos tipos de aplicações que acessam este banco de dados);
- Integridade (dependência das chaves);
- Não haver redundância de dados;
- Padronização dos dados;
- Autocontenção (armazenamento dos dados e dos metadados);
- Restauração ou reorganização (atomicidade);
- Restrições ou consistência (único local de armazenamento);
- Visões (limitação que o DBA restringe aos usuários. Segurança e privacidade).



### Saiba Mais

O(a) Administrador(a) de Banco de Dados (também conhecido como Gerente, Especialista em Banco de Dados ou pela sigla em inglês DBA), é o responsável por desenvolver e monitorar a arquitetura de dados e o banco de dados de uma empresa e/ou seus clientes.

Fonte: <http://tutano.trampos.co/16206-guia-de-profissoes-dba/>

## História do SGBD

No início, o Sistema Gerenciador de Banco de Dados (SGBD) não era desse jeito. Houve várias abordagens, que nada mais é que a maneira como os dados são organizados e como é feita a recuperação dos dados. Os bancos de dados podem ser classificados como **TRADICIONAIS** e **NÃO TRADICIONAIS**.

Nos **bancos de dados tradicionais** as informações são armazenadas e consultadas de forma numérica ou textual. As principais abordagens dos bancos de dados tradicionais – também chamadas por alguns autores de modelo – de banco de dados são: Rede, Hierárquico, Relacional e Orientado a Objetos, conforme a Tabela:

## TABELA 4 - Tipos de Abordagens ou Modelos de Banco de Dados

ABORDAGEM	Descrição
HIERÁRQUICO	Surgiu na década de 1960. Organiza os dados em uma estrutura hierárquica (árvore) em um sistema unidirecional de pai para filho, iniciando sempre pela raiz, formado por uma coleção de registros conectados por links e estes registros têm apenas um possuir (pai). Os SGBDs mais conhecidos: IMS, ADABAS e o SYSTEM 2000.
REDE	Surgiu no final da década de 1960. Organiza os dados em uma estrutura formada por várias listas que define uma rede de ligações (grafo direcionado). Os dados são organizados em tipos de registros e ligações entre dois tipos de registros. Os SGBDs mais conhecidos: IDMS, Total.
RELACIONAL	Em meados de 1960, Edgar Frank Codd (pesquisador do Laboratório de IBM em San Jose – Califórnia) pesquisava novas formas de manipular grandes volumes de dados. Uma das principais características é a possibilidade de relacionar várias tabelas evitando assim a redundância no armazenamento dos dados.
ORIENTADO A OBJETOS	Surgiu a partir da década de 1990. Possui os conceitos da orientação a objetos como encapsulamento, herança e classe. Esses meios de armazenamento se tornaram conhecidos com o crescente uso de linguagens orientadas a objeto.

INSTITUCIONAL, 2022.

## Sistema de Banco de Dados

Elmasri e Navathe (2018) explicam que o sistema de banco de dados é a união do banco de dados com o software de SGBD, conforme mostra a figura a seguir:

## FIGURA 6 - Esquema de um Sistema de Banco de Dados



HEUSER, 2011 [ADAPTADA].

Os principais softwares de banco de dados no mercado que possuem SGBD são: SQL SERVER, Oracle, MySQL, DB2, Firebird e PostgreSQL.

Já o Microsoft Access não possui todas as características, portanto, é um banco de dados, mas não é um sistema de banco de dados.

Para ser utilizado em um **ambiente de banco de dados**, o sistema de banco de dados precisa organizar em componentes que definem, regulam, coletam, armazenam e gerenciam os dados. Esse ambiente possui as seguintes partes: **HARDWARE, SOFTWARE, PESSOAS, PROCEDIMENTOS e DADOS**.

Na tabela a seguir, é apresentado um exemplo comparativo entre processamento de arquivos, SGBD hierárquico, SGBD rede, SGBD relacional e SGBD orientado a objetos:

## TABELA 5 - Exemplo de Comparativo entre as Abordagens de Dados

Descrição	Exemplo	Arquivos	Hierárquico	Rede	Relacional	Orientado a Objetos
Funcionários	Relatório contendo informações dos funcionários	ARQUIVO	TIPO DE SEGMENTO	TIPO DE REGISTRO	TABELA	CLASSE
Relatório contendo informações dos funcionários	Gertrudes Flores	REGISTRO	OCORRÊNCIA DE SEGMENTO	REGISTRO ATUAL	LINHA (TUPLA)	INSTÂNCIA DE OBJETO
Data de admissão do funcionário	04/06/2000	CAMPO	CAMPO DE SEGMENTO	CAMPO DE REGISTRO	ATRIBUTO DE TABELA	ATRIBUTO DO OBJETO
Identificação do funcionário	1234	ÍNDICE	CAMPO DE SEQUÊNCIA	CHAVE DE REGISTRO	CHAVE	IDENTIFICADOR DE OBJETO

INSTITUCIONAL, 2022.

A linguagem padrão dos bancos de dados tradicionais é a Structured Query Language (SQL), linguagem de consulta estruturada, presente nos bancos de dados do tipo relacional e orientado a objetos.

Mas com o crescimento dos dados não estruturados nas redes sociais e web, surgiu um novo conceito chamado **NoSQL**.

# NoSQL

**NoSQL** significa Not Only SQL (Não é Apenas SQL) que são sistemas para gerenciar grandes volumes de dados. Os bancos de dados NoSQL são bancos de dados não relacionais que são usados com a tecnologia Big Data e oferecem uma arquitetura muito mais escalável e eficiente que os bancos de dados relacionais para este tipo de dado.

undefined

Para aprofundar seu conhecimento, confira a leitura indicada a seguir.

## Estudo Guiado

**Leia páginas 795 a 820**

[Clique no link e leia o livro](#)

ELMASRI, R.; NAVATHE, S. B. Sistemas de banco de dados. 7. ed. São Paulo: Pearson, 2018.

## Big Data

**FIGURA 7 - Big Data**



BANCO DE IMAGENS

**Big Data**, termo usado quando há o uso de uma grande quantidade de dados com diferentes fontes e formatos. Trata-se de um conjunto de dados cujo tamanho vai além da capacidade típica das ferramentas de software de banco de dados. Atualmente possui cinco características, chamadas de 5 Vs: **Valor, Variedade, Velocidade, Veracidade e Volume** descritas na tabela a seguir:

## TABELA 6 - Características do Big Data

CARACTERÍSTICA	DESCRIÇÃO
VALOR	Os dados devem apoiar a tomada de decisão.
VARIEDADE	Grande variedade de dados proveniente de diversas fontes estruturadas ou não estruturadas.
VELOCIDADE	Rapidez como são produzidos os dados e precisam ser analisados.
VERACIDADE	Verificar se os dados coletados realmente são úteis. Há muita informação na Internet que não é verdadeira.
VOLUME	Os dados são produzidos e coletados grandes volumes de dados.

INSTITUCIONAL, 2022.

Há uma variedade de soluções atualmente para o Big Data e cresce gradativamente todo ano.

É possível encontrar o Big Data em diversas aplicações, por exemplo: carros automatizados, combate ao crime e ao terrorismo, detecção de fraudes, detecção de terremotos, detecção e prevenção de epidemias, educação customizada, por demanda e on-line, geomarketing através de smartphones, marketing personalizado, entre outros.

Para aprofundar seu conhecimento, confira a leitura indicada a seguir.

### Estudo Guiado

**Leia as páginas 821 a 861**

[Clique no link e leia o livro](#)

ELMASRI, R.; NAVATHE, S. B. Sistemas de banco de dados. 7. ed. São Paulo: Pearson, 2018.

Para conhecer mais sobre Big Data, com Dados Estruturados, Semiestruturados, Não Estruturados e Ciência de Dados, indicamos cursos gratuitos de 8 horas cada, via web, não possuem pré-requisito, em português e com emissão de certificado.



## Saiba Mais

**Curso: Big Data Fundamentos**, disponível de forma gratuita na Data Science Academy, disponível através do site: <https://www.datascienceacademy.com.br/course?courseid=big-data-fundamentos>

**Curso: Introdução à Ciência de Dados**, também disponível de forma gratuita na Data Science Academy, disponível através do site: <https://www.datascienceacademy.com.br/course?courseid=introducao--ciencia-de-dados>

## Profissões relacionadas a Banco de Dados

Na tabela a seguir, apresentamos as principais profissões relacionadas a banco de dados:

**TABELA 7 - Profissões relacionadas a Banco de Dados**

PROFISSÃO	DESCRIÇÃO
ADMINISTRADOR DE BANCO DE DADOS	Supporte técnico aos bancos de dados existentes; Personalização de bancos de dados comerciais para necessidades específicas; Planejamento e projeto de bancos de dados para necessidades específicas; Solução de problemas para atender as necessidades dos clientes; Desenvolvimento de banco de dados para uma ampla variedade de aplicações; Supervisão da instalação de novos SGBD; Criar procedimentos de backup, restauração e recuperação de desastres; Atuar com bancos de dados relacionais e não relacionais.
ANALISTA DE DADOS	Trabalhar com as equipes de TI, gestão e/ou cientistas para determinar os objetivos organizacionais; Coletar dados de fontes primárias e secundárias; Realizar limpeza nos dados e descartar informações irrelevantes; Analisar e interpretar os resultados utilizando ferramentas estatísticas e técnicas convencionais; Identificar novas oportunidades para melhoria de processos; Fornecer relatórios de dados concisos e visualizações de dados claros para a gestão; Concepção, criação e manutenção de banco de dados relacionais e NoSQL

	e sistemas de banco de dados; Resolver problemas de código e questões relacionadas a dados; Dominar linguagens (R, Python, SQL) e softwares de análise de dados (SAS, Tableau, Qlik).
ANALISTA DE NEGÓCIOS	Estabelecer os objetivos e o âmbito de sistemas de negócios e de TI; Identificar problemas organizacionais e conceber soluções orientadas a dados; Realizar análises estatísticas, pesquisas, oficinas de formação e testes; Recomendar mudanças nos processos, pessoal ou ofertas de produtos para tornar os departamentos internos mais eficientes; Conceber novos sistemas ou alterar os existentes; Fazer recomendações específicas de TI e apoiar a sua implementação; Agir como um elo de ligação entre os gestores e equipes técnicas; Propor suas decisões baseadas em dados.
ARQUITETO DE DADOS	Colaborar com as equipes de TI e gestão para elaborar uma estratégia de dados que atenda aos requisitos da empresa; Criar um inventário de dados necessários para implementar a arquitetura; Pesquisar novas oportunidades de aquisição de dados; Identificar e avaliar as atuais tecnologias de gerenciamento de dados; Criar um fluxo de dados dentro da empresa; Desenvolver modelos de dados; Projetar, documentar, construir e implementar arquiteturas e aplicações de banco de dados, por exemplo, grandes bancos de dados relacionais e NoSQL; Integrar a funcionalidade técnica, por exemplo: escalabilidade, segurança, desempenho, recuperação de dados, confiabilidade etc. Implementar medidas para assegurar a precisão dos dados e acessibilidade; Monitorar constantemente, aperfeiçoar e apresentar um relatório sobre o desempenho dos sistemas de gerenciamento de banco de dados.
CIENTISTA DE DADOS	Comunicar previsões e resultados para a gestão e os departamentos de TI através de visualizações de dados eficazes; Extrair grandes volumes de dados de múltiplas fontes internas e externas; Empregar os programas de análise sofisticadas, aprendizado de máquina e métodos estatísticos para preparar os dados para uso em modelagem preditiva e prescritiva; Explorar e analisar dados de uma variedade de ângulos para determinar fraquezas escondidas, tendências e/ou oportunidades; Conceder soluções orientadas a dados para os principais desafios da empresa; Criar algoritmos para resolver problemas e criar ferramentas para automatizar o trabalho; Recomendar mudanças econômicas aos procedimentos e estratégias existentes; Dominar técnicas de análise e armazenamento de dados.
ENGENHEIRO DE DADOS	Responsável pela arquitetura e infraestrutura de dados que trabalha com design, construção e manutenção de sistemas para armazenamento e processamento de dados.
PROJETISTA OU ADMINISTRADOR DE BANCO DE DADOS	Identificação das necessidades de informações da organização; Retirar da equipe de desenvolvimento a responsabilidade da organização e estruturação dos dados, trabalhando na análise e descrição geral de dados, na definição do modelo conceitual, no projeto lógico do banco de dados e, ainda, na análise funcional dos dados; Garantia da validade, exatidão, consistência e da disponibilidade dos dados; Promoção do compartilhamento dos dados da organização; Disponibilidade de informações suficientes que facilitem o uso e a compreensão dos dados; Refinamento das informações da organização visando melhorar a sua qualidade e performance de acesso; Impedimento do crescimento desordenado de objetos da base de dados; Tomada de medidas preventivas e corretivas para identificação de problemas existentes no modelo de dados da organização;

Fornecimento de suporte na compreensão e utilização dos dados à equipe de desenvolvimento; Manutenção dos modelos de informações atualizados.

INSTITUCIONAL, 2022.



## Saiba Mais

### SALÁRIOS NA ÁREA DE TECNOLOGIA

O Guia Salarial da Robert Half é uma das mais respeitadas fontes de informação sobre remuneração e tendências de recrutamento para auxiliar empresas e profissionais a tomarem as melhores decisões e publicada anualmente um guia salarial disponível gratuitamente em: <https://www.roberthalf.com.br/guia-salarial>

## Projeto de Banco de Dados

Chamamos de projetista de banco de dados ou administrador de dados, o responsável por criar o projeto de banco de dados.

Um projetista de banco de dados define toda a estrutura de armazenamento e elementos necessários para o banco de dados.

Um projetista deve ter conhecimento sobre:

- Análise orientada a objetos;
- Avaliar as necessidades dos usuários;
- Entendimento da linguagem e do ambiente de desenvolvimento;
- Modelagem de banco de dados;
- Técnicas de projeto de banco de dados.

Todo bom sistema de banco de dados deve apresentar um projeto, que visa à organização das informações e utilização de técnicas para que o futuro sistema obtenha boa performance e também facilite infinitamente as manutenções que venham a acontecer. Um sistema de informação utiliza um banco de dados que necessita de **DADOS** e **PROCESSOS**, onde

## DADOS

Parte mais estável do sistema, após definido pouco se modifica (objeto desta Unidade Curricular);

## PROCESSOS

Parte mais dinâmica, pois varia de acordo com as necessidades do negócio.

O **PROJETO DE BANCO DE DADOS** tem o objetivo de transformar as necessidades de informações no negócio em um banco de dados. Ou seja, foca na maneira em que o banco de dados será usado para armazenar e gerenciar os dados do usuário final e a primeira etapa é criar o modelo de dados.

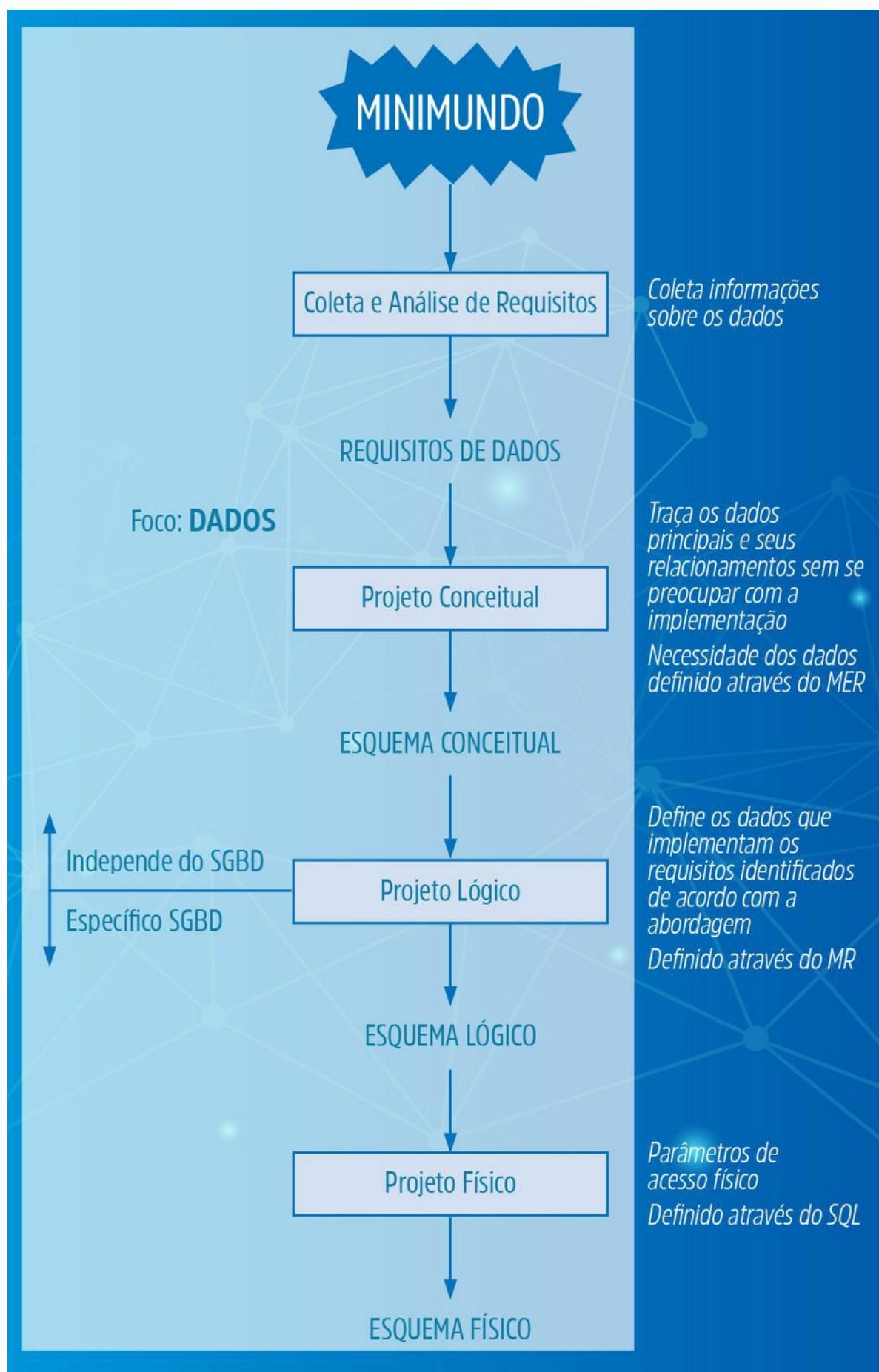
O projeto de um novo banco de dados dá-se em três fases: **Modelo Conceitual, Modelo Lógico e Modelo Físico de Dados**. As duas primeiras fases (Modelo Conceitual e Dados e Modelo Lógico de Dados) são níveis de abstração na construção do modelo de dados (HEUSER, 2011).

O projeto de um novo banco de dados começa em uma fase chamada especificação e análise de requisitos (ELMASRI; NAVATHE, 2018).

Esses requisitos são documentados com detalhes e transformados em um **projeto conceitual**, que pode ser representado e manipulado com o uso de algumas ferramentas computadorizadas, de modo que possa ser facilmente mantido, modificado e transformado em uma implementação de banco de dados. Esse projeto é então traduzido para um projeto lógico que pode ser expresso em um modelo de dados implementado em um sistema gerenciador de banco de dados (SGBD). A última etapa é o **projeto físico**, durante o qual outras especificações são fornecidas para armazenar e acessar o banco de dados. Então o projeto de banco de dados é implementado, preenchido com dados reais e mantido para refletir o estado do minimundo. (ELMASRI; NAVATHE, 2018).

A figura a seguir esquematiza as etapas do projeto de banco de dados.

**FIGURA 8 - Etapas do Projeto de Banco de Dados**



A definição ou informação descritiva do banco de dados também é armazenada no SGBD na forma de um catálogo ou dicionário, chamado de **metadados**. Um banco de dados representa algum aspecto do mundo real, às vezes chamado de **minimundo** ou de **universo de discurso (UoD)**. As mudanças no minimundo são refletidas no banco de dados (ELMASRI; NAVATHE, 2018).

Vamos conhecer cada etapa deste projeto e suas características, na tabela a seguir:

**TABELA 8 - Etapas do Projeto de Banco de Dados**

ETAPA	DESCRÍÇÃO
<b>PROJETO CONCEITUAL</b>	<p>Modelo abstrato de banco de dados de acordo com a visão da empresa.</p> <p>Agrupa os elementos de dados no banco de dados.</p> <p>Identifica elementos redundantes e os grupos de elementos de dados repetidos para aplicações específicas.</p> <p><b>É uma das etapas mais importantes do projeto.</b></p> <p>Cria-se o Modelo Entidade Relacionamento (MER).</p>
<b>PROJETO LÓGICO</b>	<p>Mapeia o modelo conceitual para um modelo lógico.</p> <p>Padroniza de acordo com a implementação, adequando ao tipo de abordagem de SGBD.</p>
<b>PROJETO FÍSICO</b>	Criação dos objetos no software de banco de dados através de scripts, usando a linguagem SQL (Structured Query Language - Linguagem de Consulta Estruturada).

INSTITUCIONAL, 2022.



### ATENÇÃO

Há ainda outra forma de mostrar as informações do projeto chamado de **ESQUEMA RELACIONAL** que é uma representação textual de uma tabela (também chamado de **relação**) na abordagem relacional, onde os atributos aparecem entre parênteses e os atributos-chave aparecem sublinhados ou negrito. Ex.: ESTUDANTE (**matrícula**, nome, data\_de\_nascimento, estado\_civil).

# Modelo de Dados

Um modelo de dados é descrição formal da estrutura de um banco de dados (HEUSER, 2011).

Modelo de (banco de) dados é uma descrição dos tipos de informações que estão armazenadas em um banco de dados. Por exemplo, no caso da indústria acima citado, o modelo de dados poderia informar que o banco de dados armazena informações sobre produtos e que, para cada produto, são armazenados seu código, preço e descrição. Observe que o modelo de dados não informa quais os produtos que estão armazenados no banco de dados, mas apenas que o banco de dados contém informações sobre produtos. (HEUSER, 2011).

O conjunto de conceitos usados na construção de um modelo dá-se o nome de **ABORDAGEM DE MODELAGEM** (HEUSER, 2011).

## Os benefícios da modelagem de dados são:

- Entendimento do Negócio;
- Comunicação;
- Estruturação dos Dados;
- Definição de Escopo;
- Estimativas de prazos e custos.

Para aprofundar seu conhecimento, confira a leitura indicada a seguir.

## Estudo Guiado

**Leia páginas 19 a 29 do livro Projeto de Banco de Dados - Capítulo 1 - Introdução.**

[Clique no link e leia o livro](#)

HEUSER, C. A. Projeto de banco de dados. 6. ed. Porto Alegre: Bookman, 2011.

## Exemplos das etapas do projeto de Banco de Dados

O estudo de caso apresentado tem o objetivo de apresentar as etapas do projeto de banco de dados:

### Estudo de Caso

#### CLÍNICA MÉDICA

Uma especialidade médica possui vários médicos, por exemplo, clínica geral, pediatria entre outros. Há médicos, que podem ter mais de uma especialidade, por exemplo, um médico que é alergologista e também homeopata. Para isso, como dados a serem armazenados temos a sigla e o nome da especialidade médica e dos médicos temos o CRM, o nome e o telefone para contato.

# Modelo conceitual de Dados

Nesta primeira etapa, criamos o Modelo Entidade e Relacionamento (M.E.R.) que demonstra as entidades com seus atributos (dados) e os relacionamentos entre essas entidades. A figura a seguir mostra a resolução do estudo de caso no modelo conceitual de dados:

## FIGURA 9 - Modelo Conceitual de Dados

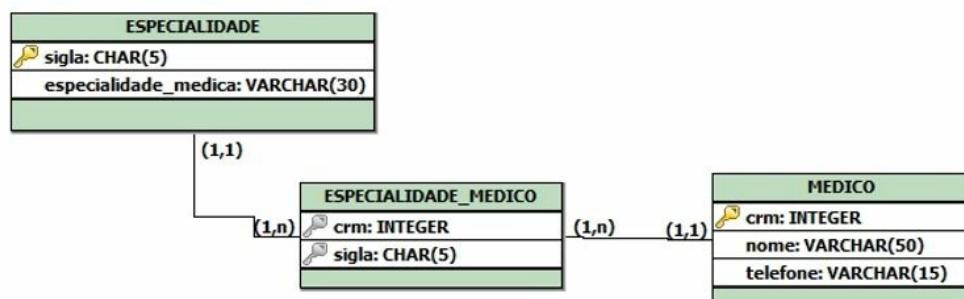


INSTITUCIONAL, 2022.

## Modelo lógico de Dados

Nesta etapa, criamos o Modelo Lógico de Dados que demonstra as entidades com seus atributos (dados) e os relacionamentos entre essas entidades. A figura a seguir mostra a resolução do estudo de caso no modelo lógico de dados, adequado para a abordagem relacional com os tipos de dados da linguagem SQL:

**FIGURA 10 - Exemplo Modelo Lógico de Dados**

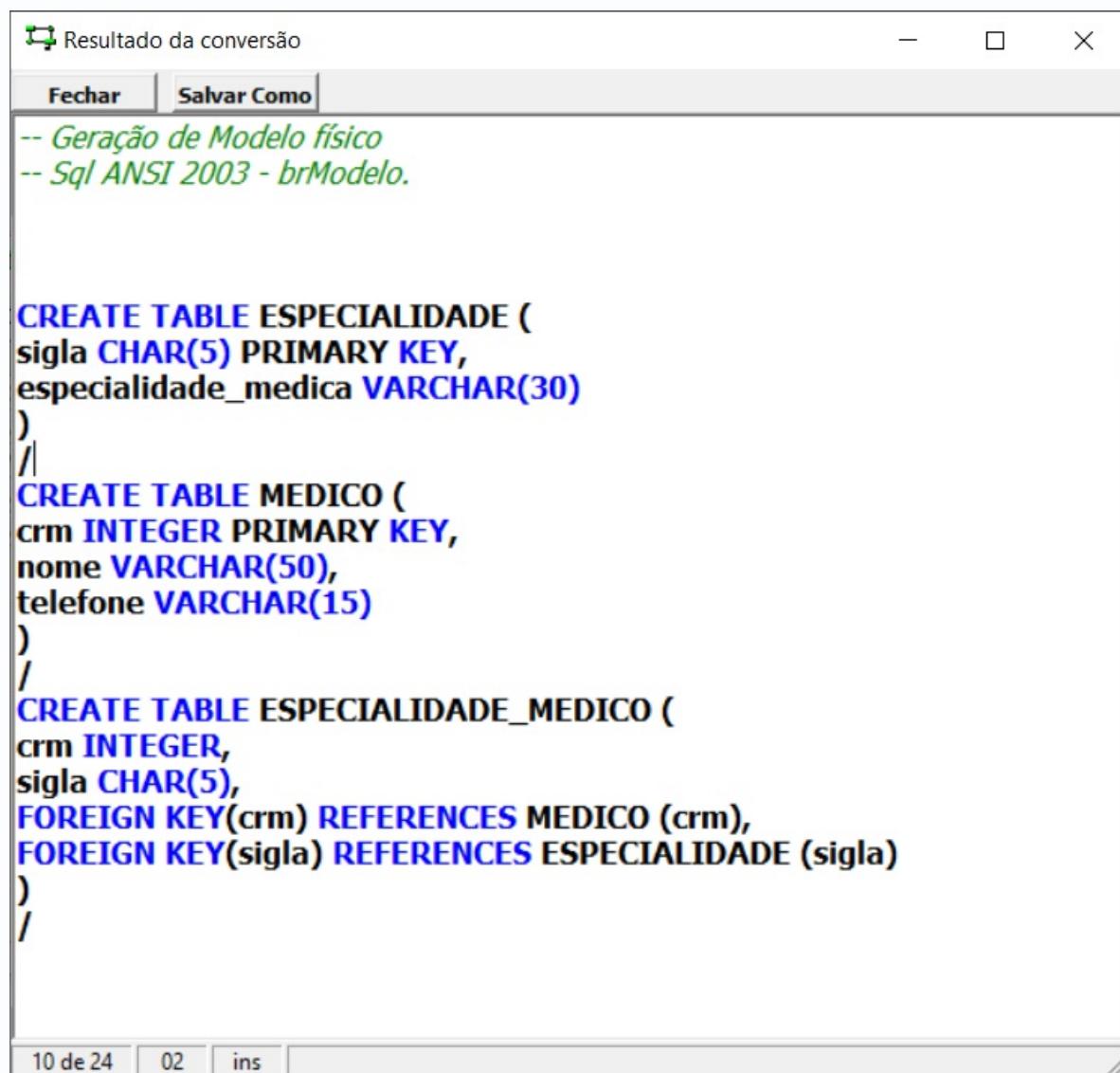


INSTITUCIONAL, 2022.

## Modelo físico de Dados

Na última etapa, é criado o Modelo Físico de Dados, onde é gerado o script (código contendo os comandos) da criação das tabelas com os campos e relacionamentos, baseado na linguagem SQL (Structured Query Language). A Figura exemplifica o script contendo a criação de tabelas com restrições de chave primária e chave estrangeira:

## FIGURA 11 - Modelo Físico de Dados



The screenshot shows a software interface with a title bar 'Resultado da conversão'. Below the title bar are two buttons: 'Fechar' and 'Salvar Como'. The main area contains the generated SQL code:

```
-- Geração de Modelo físico
-- Sql ANSI 2003 - brModelo.

CREATE TABLE ESPECIALIDADE (
    sigla CHAR(5) PRIMARY KEY,
    especialidade_medica VARCHAR(30)
)
/
CREATE TABLE MEDICO (
    crm INTEGER PRIMARY KEY,
    nome VARCHAR(50),
    telefone VARCHAR(15)
)
/
CREATE TABLE ESPECIALIDADE_MEDICO (
    crm INTEGER,
    sigla CHAR(5),
    FOREIGN KEY(crm) REFERENCES MEDICO (crm),
    FOREIGN KEY(sigla) REFERENCES ESPECIALIDADE (sigla)
)
```

At the bottom of the window, there is a navigation bar with the text '10 de 24' and '02 ins'.

INSTITUCIONAL, 2022.

## Esquema relacional

É possível representar o esquema relacional das tabelas utilizando a notação de relação, onde os atributos-chave são diferenciados por estar em negrito ou sublinhado, conforme exemplo abaixo:

undefined

Para aprofundar seu conhecimento, confira a leitura indicada a seguir.

## Estudo Guiado

**Capítulo 4 - Modelagem de dados - Estágios da modelagem de dados, páginas 77 a 80.**

[Clique no link e leia o livro](#)

PUGA, S.; FRANÇA, E.; GOYA, M. Banco de dados: implementação em SQL, PL/SQL e Oracle 11G. São Paulo: Pearson, 2013.

Chegamos ao fim desta unidade, confira a seguir, uma breve síntese sobre o que vimos até aqui.

# Recurso Externo

Recurso é melhor visualizado no formato interativo

# Referências

## Bibliográficas

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD).

**Diário Oficial da União.** Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2018/Lei/L13709compilado.htm](http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/Lei/L13709compilado.htm). Acesso em: 1º jun. 2020.

ELMASRI, R.; NAVATHE, S. B. **Sistemas de banco de dados.** 7. ed. São Paulo: Pearson, 2018.

HEUSER, C. A. **Projeto de banco de dados.** 6. ed. Porto Alegre: Bookman, 2011.

INFOGRAPHIC: what happens in an internet minute 2020. **Allaccess**, mar. 2020. Disponível em: <https://www.allaccess.com/merge/archive/31294/infographic-what-happens-in-an-internet-minute>.

Acesso em: 26 jun. 2020.

ORACLE ACADEMY. [2020]. Disponível em: <https://academy.oracle.com/pt-br/oa-institution.html>.  
Acesso em: 6 jun. 2020.

PUGA, S.; FRANÇA, E.; GOYA, M. **Banco de dados:** implementação em SQL, PL/SQL e Oracle 11G. São Paulo: Pearson, 2013.

RIPARI, C. Por que os dados são considerados o novo petróleo. **It Forum 365**, mar. 2019. Disponível em: <https://itforum365.com.br/por-que-dados-sao-considerados-o-novo-petroleo/>. Acesso em: 1º jun. 2020.