



**Digital
College**

ENSINO DE HABILIDADES DIGITAIS

Formação em Data Analytics

Aula 01

Facilitadora: **Alex Souza**



Quem sou?

Alex Souza





DA15 - Data Analytics

Grupo do WhatsApp



Leia o QR code com a câmera do WhatsApp
para entrar neste grupo



Sobre o Curso

Unidade 01 – Banco de Dados

Unidade 02 – ETL

Unidade 03 – Python para Análise de Dados

Unidade 04 – Power BI Desktop



Sobre o Curso

Unidade 01 – Banco de Dados

- 1 – Banco de Dados Relacionais
- 2 – Banco de Dados Dimensionais
- 3 – SQL Avançado

Unidade 02 – Extração, Transformação e Carga - ETL

- 1 - Business Intelligence
- 2 – Extração, Transformação e Carga de Dados
- 3 -Pipeline de Dados

Unidade 03 – Python para Análise de Dados

- 1 – Iniciando com Python
- 2 – Manipulação de Dados com Python
- 3 – Pandas para Análise de Dados
- 4 – Gráficos com Python

Unidade 04 – Power BI Desktop

- 1 – Conhecendo o Power BI
- 2 – KPI's e Métricas
- 3 – Modelagem, Relacionamentos e DAX
- 4 – Visualização de Dados



Ferramentas e Tecnologias

- Arquivos – xls, txt, json, pdf, xml, etc
- Modelagem – Dbesquema, BRModel
- Banco de Dados – PostgreSQL - PgAdmin
- Python Anaconda – Jupyter Notebook e VS Code
- Power BI
- Github.



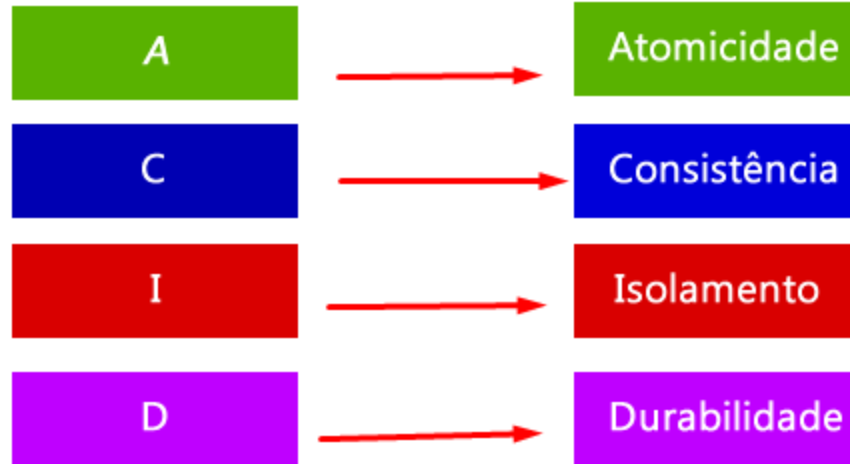
Banco de Dados

Um ***Sistema de Gerenciamento de Banco de Dados (SGBD)*** consiste em uma coleção de dados inter-relacionados e em um conjunto de programas para acessá-los.

Um ***Banco de Dados*** é um conjunto de dados inter-relacionados.



Propriedades de um Banco de Dados Relacional



Atomicidade

Seria a garantia de que a transação será feita totalmente ou não será feita. Nesse caso, a transação não é feita "pela metade". Se por ventura uma operação da transação falhar, conseqüentemente, toda a transação falhará.

Consistência

Seria a proteção da integridade dos dados. Ou seja, se um banco de dados fizer uma operação que não seja válida, o processo será impedido e retornará para o estado inicial do processo.



Isolamento

A capacidade de isolamento seria o fato de uma transação não “atrapalhar” a outra e ocorrer de forma isolada, garantindo que sejam feitas de forma individual.

Durabilidade

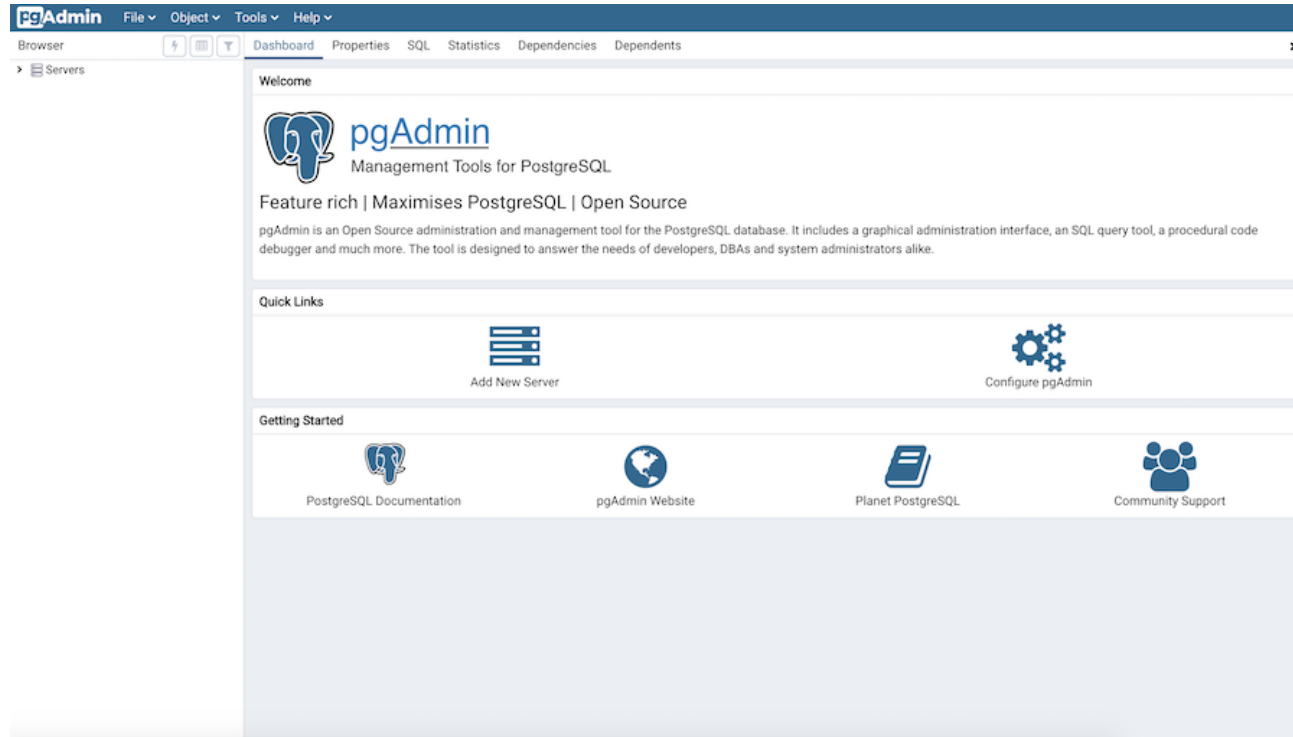
Seria a preservação dos dados após as operações terem sido realizadas. Ou seja, uma vez que uma transação for efetuada, ela permanecerá dessa forma, mesmo que ocorram problemas graves no sistema, sem precisar de retrabalho.



Abstração dos Dados

- **Nível Físico** – o nível mais baixo de abstração descreve *como* os dados estão realmente armazenados. Num nível físico, complexas estruturas de dados de baixo nível são descritas em detalhe.
- **Nível Conceitual** – este nível descreve *quais* dados estão armazenados de fato no banco de dados e as relações que existem entre eles. O nível conceitual é usado por administradores do banco de dados, que podem decidir quais informações devem ser mantidas no banco de dados.
- **Nível de Visões** – a maioria dos usuários do sistema de banco de dados não está interessada em todas as informações existentes no banco de dados. Cada grupo de usuários deve enxergar somente os dados que lhe dizem respeito. Assim, cada grupo de usuários tem uma visão do banco de dados. O nível mais alto de abstração é composto de visões que cada grupo de usuários tem do banco de dados.

Postgres / Pgadmin



O que é Banco de Dados?

Conjunto de dados que se relacionam.

- **Propriedades:**

- coleção lógica e coerente de dados com um significado inerente;
- projetado, construído e populado com dados para um propósito específico;
- possui um conjunto pré definido de usuários e aplicações;
- representa algum aspecto do mundo real - “mini-mundo” e qualquer alteração efetuada neste mini-mundo é automaticamente refletida no banco de dados.

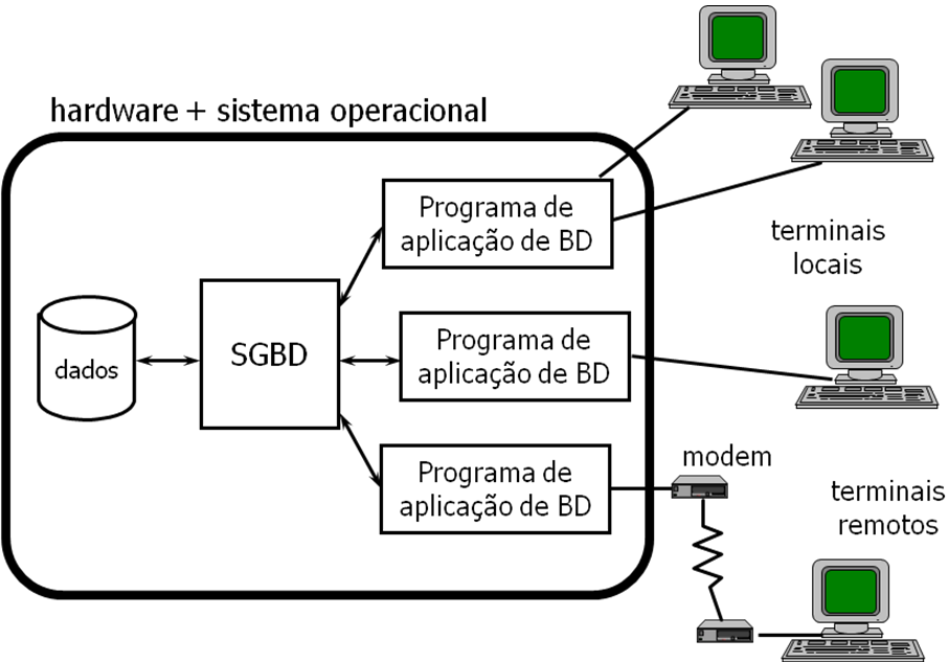
Benefícios de uso de um SGBD

- Implementação do Conceito de transação
- Controle de Concorrência
- Mecanismo de acesso aos dados
- Mecanismo de recuperação (Tolerância a falhas)
- Maior Segurança

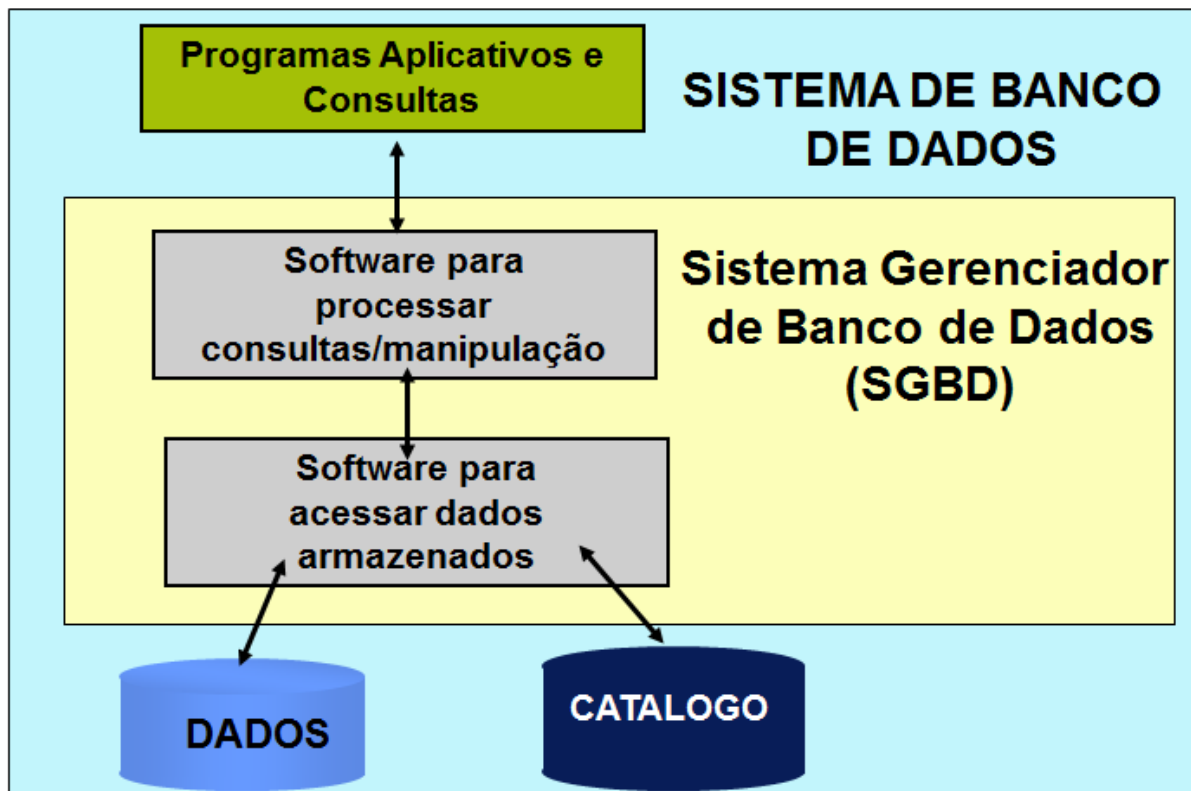


Evolução dos Bancos de Dados

Bancos de Dados Centralizados



Estrutura Geral de um Sistema de Banco de Dados



Instância

É o **conjunto de ocorrências** dos objetos de dados de um esquema em um dado momento do tempo (**Estado do BD num momento do tempo**).

	PESSOA			DEPART	
CPF	NOME	COD_DEP	COD_DEP	NOME	LOCAL
123	A	D-10	D-10	DEPART-10	L1
456	B	D-10	D-20	DEPART-20	L2
789	C	D-20	D-30	DEPART-30	L3

ESQUEMA

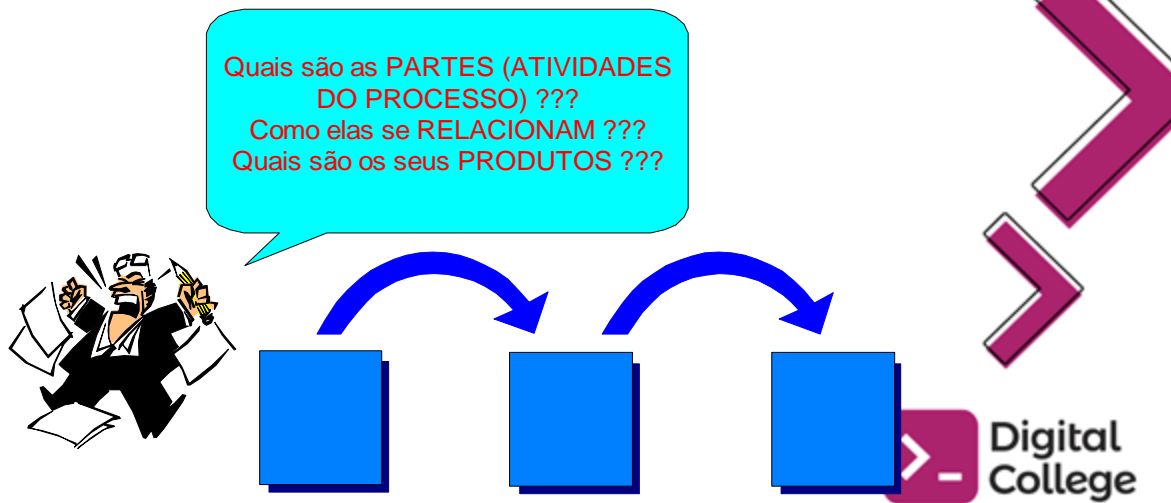
INSTÂNCIA



Projeto de Banco de Dados

É o processo de:

- **determinar a organização** de um banco de dados, incluindo sua estrutura, conteúdo e aplicações;
- **desenvolver a estrutura** de um banco de dados;
- **projetar a estrutura lógica e física** de um ou mais bancos de dados com a finalidade de acomodar as informações necessárias aos usuários de uma organização, para um definido conjunto de aplicações.



Projeto de Banco de Dados

Fase 1 – **Projeto Conceitual** de Banco de Dados:

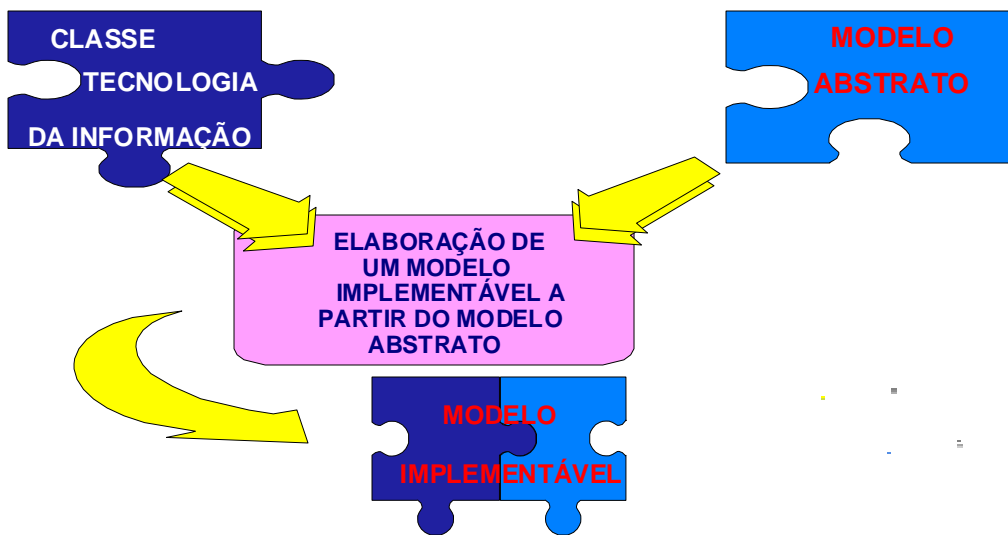
Elaboração de um modelo abstrato do mini mundo – totalmente independente de qualquer tecnologia da informação.



Projeto de Banco de Dados

Fase 2 – Projeto Lógico de Banco de Dados

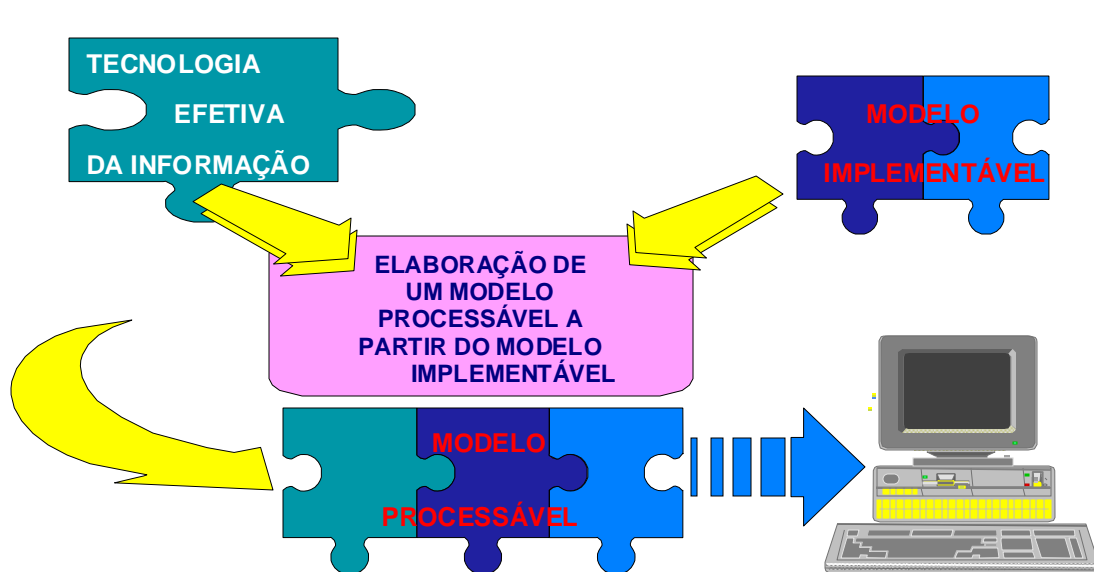
Elaboração de um modelo implementável durante a qual endereçamos uma classe de tecnologia da informação (Ex: Bancos de Dados Relacionais).



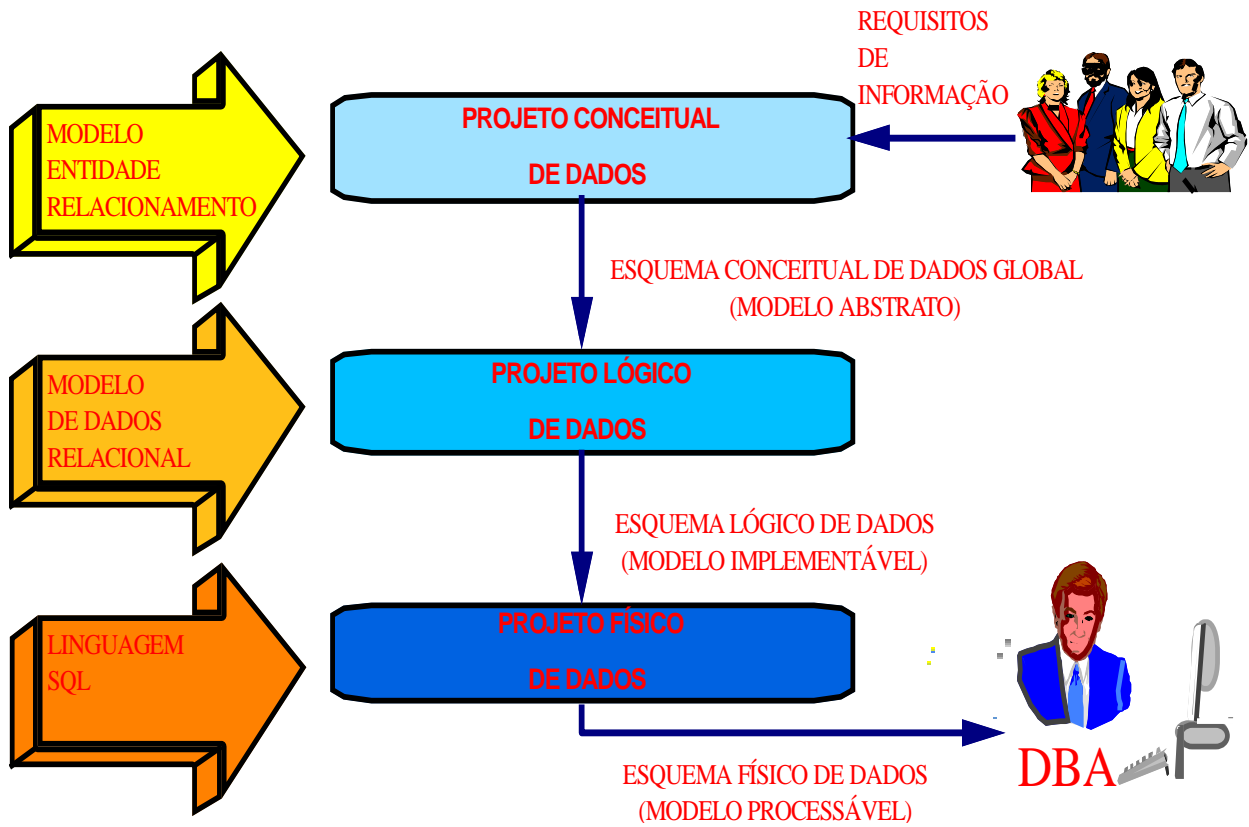
Projeto de Banco de Dados

Fase 3 – Projeto Físico de Banco de Dados

Elaboração de um modelo processável, durante a qual endereçamos uma **tecnologia da informação específica** (Ex: Bancos de Dados Relacional ORACLE versão 11g).

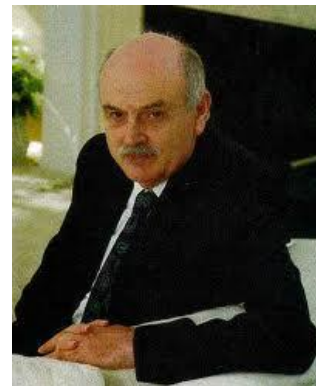


Ferramentas de Projeto



Modelo Relacional

- Foi proposto por Edgar Codd em 1970
- Modelo Dominante nos SGBD Comerciais
- Possui uma extensão denominada Objeto Relacional



Componente do Modelo Relacional

Relação = Tabela

TABELA DEPARTAMENTO		
ID	NOME	ID_REGIAO
10	VENDAS	1
20	OPERAÇÕES	1
30	ADMINISTRATIVO	2
40	FINANCEIRO	3

TABELA REGIÃO	
ID	NOME
1	NORTE
2	CENTRO
3	SUL



Componente do Modelo Relacional

Atributo = Coluna

TABELA DEPARTAMENTO		
ID	NOME	ID_REGIAO
10	VENDAS	1
20	OPERAÇÕES	1
30	ADMINISTRATIVO	2
40	FINANCEIRO	3

Atributos



Componente do Modelo Relacional

Valores dos Atributos = Domínio

TABELA DEPARTAMENTO		
ID	NOME	ID_REGIAO
10	VENDAS	1
20	OPERAÇÕES	1
30	ADMINISTRATIVO	2
40	FINANCEIRO	3

Domínio



Componente do Modelo Relacional

Conjunto de Atributo = Esquema

TABELA DEPARTAMENTO		
ID	NOME	ID_REGIAO
10	VENDAS	1
20	OPERAÇÕES	1
30	ADMINISTRATIVO	2
40	FINANCEIRO	3

← Esquema



Componente do Modelo Relacional

Tupla = Linha

TABELA DEPARTAMENTO		
ID	NOME	ID_REGIAO
10	VENDAS	1
20	OPERAÇÕES	1
30	ADMINISTRATIVO	2
40	FINANCEIRO	3

← Tupla



Componente do Modelo Relacional

Conjunto de Tuplas = Instância

TABELA DEPARTAMENTO		
ID	NOME	ID_REGIAO
10	VENDAS	1
20	OPERAÇÕES	1
30	ADMINISTRATIVO	2
40	FINANCEIRO	3

← Instância



Componente do Modelo Relacional

Campo = Encontro de Linha e Coluna

TABELA DEPARTAMENTO		
ID	NOME	ID_REGIAO
10	VENDAS	1
20	OPERAÇÕES	1
30	ADMINISTRATIVO	2
40	FINANCEIRO	3

Campo



Componente do Modelo Relacional

Conteúdo do Campo = Valor

TABELA DEPARTAMENTO		
ID	NOME	ID_REGIAO
10	VENDAS	1
20	OPERAÇÕES	1
30	ADMINISTRATIVO	2
40	FINANCEIRO	3

Valor

Campo



Componente do Modelo Relacional

Campos sem Conteúdo = Nulo

Dependentes				
Responsável	Id	Nome	Grau	CPF
101	1	Pedro Antunes	Filho	12547896521
101	2	Paula Antunes	Esposa	23654789541
102	1	Marcos Leitão	Marido	12457896231
103	1	Carlos Ventura	Filha	

Nulo – campo não possui valor



Componente do Modelo Relacional

Chave Primária

TABELA DEPARTAMENTO		
ID	NOME	ID_REGIAO
10	VENDAS	1
20	OPERAÇÕES	1
30	ADMINISTRATIVO	2
40	FINANCEIRO	3

TABELA REGIÃO	
ID	NOME
1	NORTE
2	CENTRO
3	SUL

Chave Primária (PK)



Componente do Modelo Relacional

Chave Primária = Atributo Único

TABELA DEPARTAMENTO		
ID	NOME	ID_REGIAO
10	VENDAS	1
20	OPERAÇÕES	1
30	ADMINISTRATIVO	2
40	FINANCEIRO	3

TABELA REGIÃO	
ID	NOME
1	NORTE
2	CENTRO
3	SUL

Chave Primária (PK)

Integridade de Chave Primária



Componente do Modelo Relacional

Chave Primária = Atributo Único

TABELA DEPARTAMENTO		
ID	NOME	ID_REGIAO
10	VENDAS	1
20	OPERAÇÕES	1
20	ADMINISTRATIVO	2
40	FINANCEIRO	3

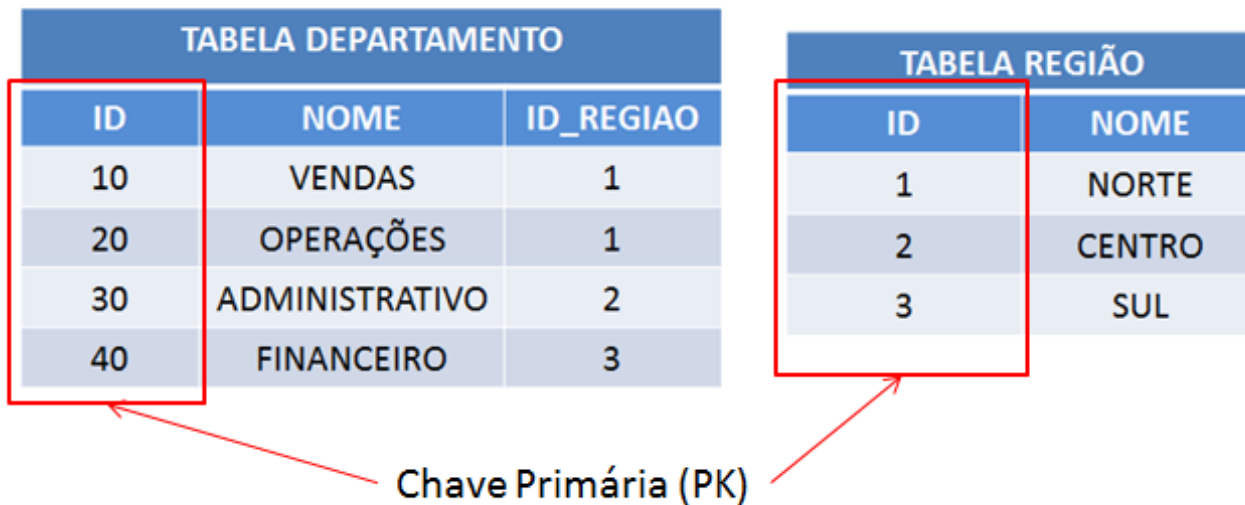
PK

Integridade de Chave Primária



Componente do Modelo Relacional

Chave Primária = Atributo Obrigatório



Integridade de Entidade



Componente do Modelo Relacional

Chave Primaria = Atributo Obrigatório

TABELA REGIÃO	
ID	NOME
	NORTE
2	CENTRO
	SUL

Nulo

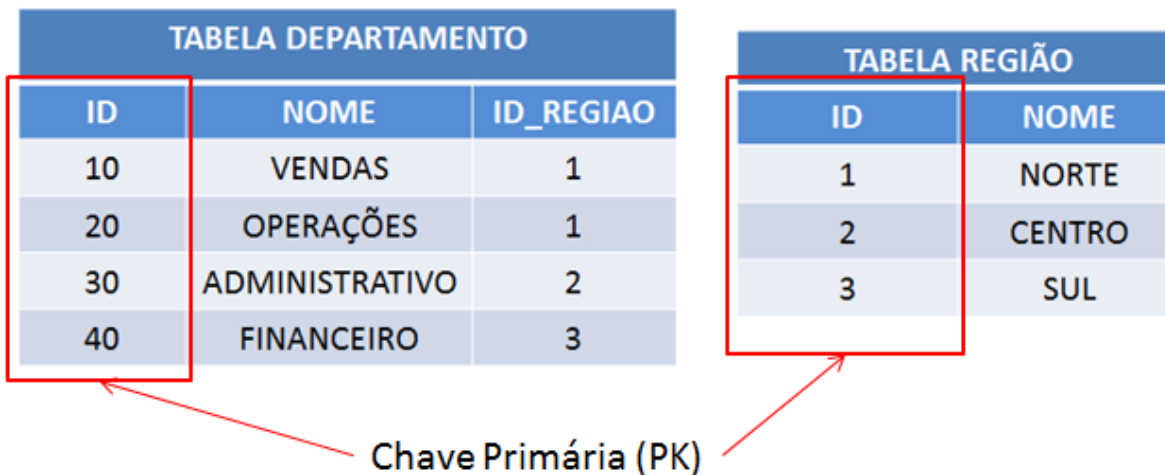
PK

Integridade de Entidade



Componente do Modelo Relacional

Chave Primária Simples



Componente do Modelo Relacional

Chave Primária Composta

Dependentes				
Responsável	Id	Nome	Grau	CPF
101	1	Pedro Antunes	Filho	12547896521
101	2	Paula Antunes	Esposa	23654789541
102	1	Marcos Leitão	Marido	12457896231
103	1	Carlos Ventura	Filho	

Chave Primária Composta



Componente do Modelo Relacional

Chave Candidata

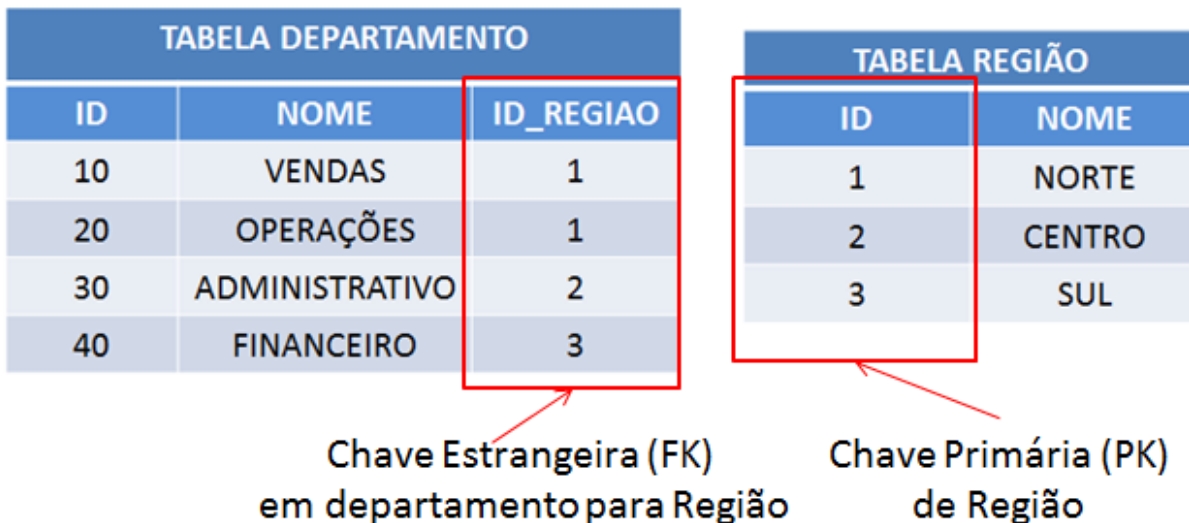
TABELA DEPARTAMENTO		
ID	NOME	ID_REGIAO
10	VENDAS	1
20	OPERAÇÕES	1
30	ADMINISTRATIVO	2
40	FINANCEIRO	3

Chaves Candidatas



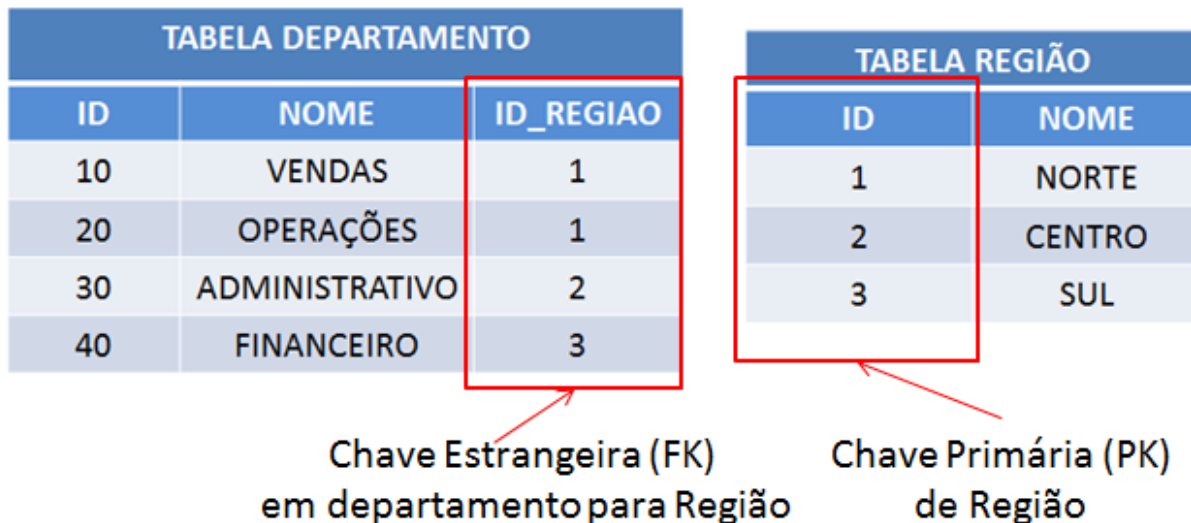
Componente do Modelo Relacional

Chave Estrangeira



Componente do Modelo Relacional

Chave Estrangeira



Chave Estrangeira -> Integridade Referencial

Vamos Praticar?

1. Criar o seguinte banco de dados: biblioteca

```
CREATE DATABASE biblioteca;
```

2. Criar tabela livros

```
CREATE TABLE IF NOT EXISTS public.livros  
(  
    id serial NOT NULL,  
    titulo character varying(100),  
    autor character varying(100),  
    ano_publicacao integer,  
    genero character varying ,  
    CONSTRAINT livros_pkey PRIMARY KEY (id)  
)
```



Vamos Praticar?

3. Inserir valores na tabela livros

```
INSERT INTO public.livros (titulo, autor, ano_publicacao, genero)  
VALUES ('Dom Quixote', 'Miguel de Cervantes', 1605, 'Romance');
```

```
INSERT INTO public.livros (titulo, autor, ano_publicacao, genero)  
VALUES ('1984', 'George Orwell', 1949, 'Ficção Distópica');
```

```
INSERT INTO public.livros (titulo, autor, ano_publicacao, genero)  
VALUES ('O Senhor dos Anéis', 'J.R.R. Tolkien', 1954, 'Fantasia');
```

```
INSERT INTO public.livros (titulo, autor, ano_publicacao, genero)  
VALUES ('Crime e Castigo', 'Fiódor Dostoiévski', 1866, 'Romance');
```

```
INSERT INTO public.livros (titulo, autor, ano_publicacao, genero)  
VALUES ('Orgulho e Preconceito', 'Jane Austen', 1813, 'Romance');
```



FIM

