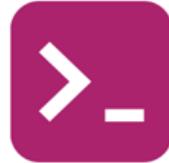




Digital
College

ENSINO DE HABILIDADES DIGITAIS

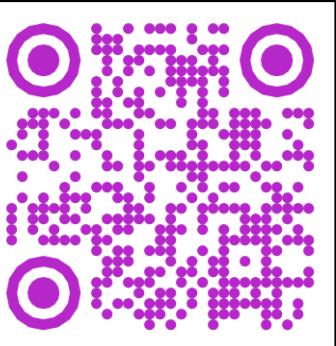


Digital
College

Formação em Data Analytics

Aula 01

Professor: **MSc. Alex Souza**



Sobre vocês

- Nome
- Formação
- Profissão
- Experiência com dados?
- Qual seu objetivo com o curso?

Formulário de alinhamento:



Sobre o curso (Unidades)

Unidade 01 – Banco de Dados

Unidade 02 – ETL

Unidade 03 – Python para Análise de Dados

Unidade 04 – Power BI Desktop



Sobre o curso (Unidades e Módulos)

Unidade 01 – Banco de Dados

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



Sobre o curso (Unidades e Módulos)

Unidade 02 – Extração, Transformação e Carga -
ETL

- 1 – Processo de Descoberta de Conhecimento
- 2 – Business Intelligence
- 3 – Extração, Transformação e Carga de Dados
- 4 – Pipeline de Dados



Sobre o curso (Unidades e Módulos)

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



Sobre o curso (Unidades e Módulos)

Unidade 04 – Power BI Desktop

1 – Conhecendo o Power BI

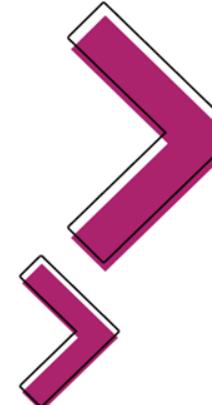
2 – Modelagem, Relacionamentos e DAX

3 – Visualização de Dados



Dicas iniciais

- Não armazenem arquivos localmente, utilizem alguma plataforma on-line (google drive, onedrive...)
- Evite utilizar pen drive ou HD externo
- Não é permitido consumo de comida na sala
- Ao sair fechem todos os programas utilizados e baixe a tampa do notebook



Ferramentas e Tecnologias

- Arquivos – xls, txt, json, pdf, xml, etc..
- BrModelo Web ([link](#))
- Banco de Dados – PostgreSQL - PgAdmin
- Python Anaconda – Jupyter Notebook e VS Code
- Power BI
- Github, Trello, Kaggle, etc..
- Ferramentas Case



Github



Todo material de aula, scripts e dicas ficarão em um repositório do github

- Criar uma conta no github
- Me enviar o nome do usuário (para eu compartilhar o repositório)

Repositório

- https://github.com/aasouzaconsult/DC_Data-Analytics_01



Contato

O que preferem?

- Discord 
- Whatsapp 







DC - Análise de Dados - 01

Grupo do WhatsApp



Escaneie o código QR com a câmera do WhatsApp para entrar neste grupo.

Unidade 1 – Banco de Dados

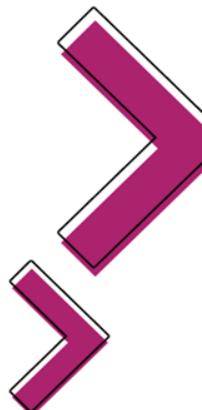


Habilidades Comportamentais

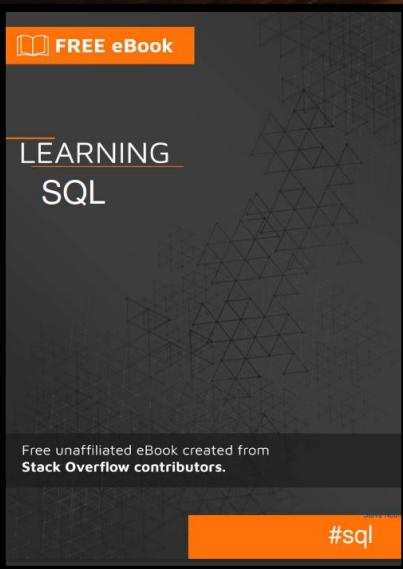
- Pensamento crítico
- Resolução de problemas
- Comunicação efetiva
- Escuta Ativa
- Habilidade de apresentação
- Networking e habilidade para trabalhar em equipe
- Inteligência Emocional e Empatia
- Aprendizado contínuo (*lifelong learner*)
- Adaptabilidade (ferramenta, negócio...)



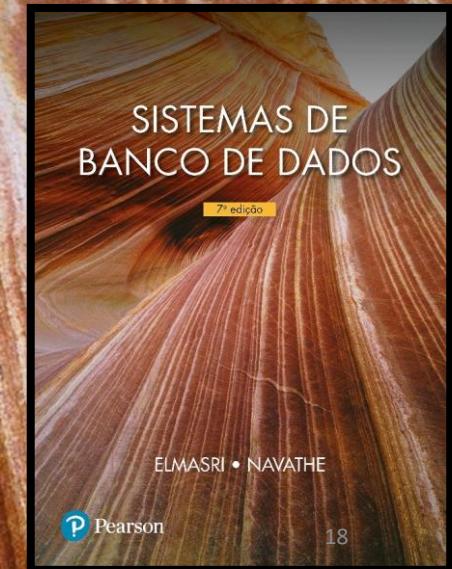
Módulo 1: Bancos de Dados Relacionais



SISTEMAS DE BANCO DE DADOS



7^a edição



Excel para Análise de Dados

Com o Excel, é possível importar, manipular, analisar e visualizar dados de diferentes fontes, como bancos de dados, arquivos de texto, arquivos CSV e outros.



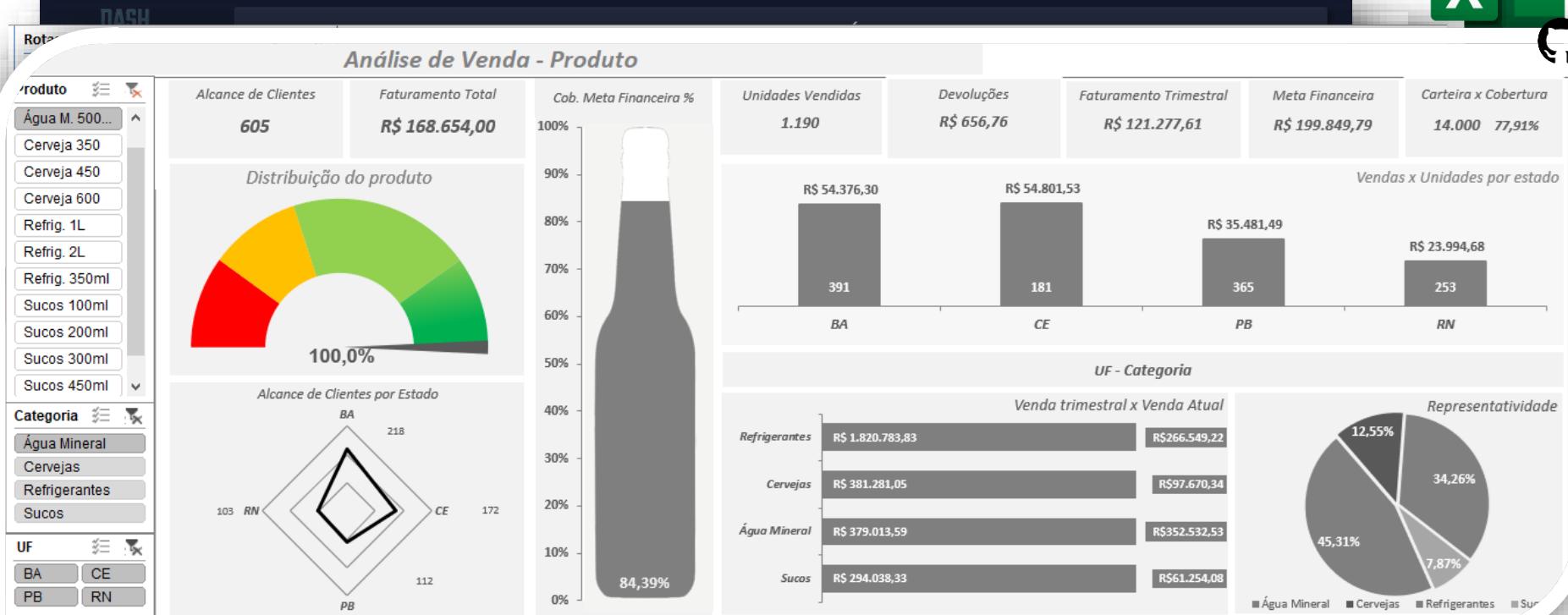
Algumas das principais funções do Excel para análise de dados incluem:

- Classificação e filtragem de dados
- Criação de tabelas dinâmicas e gráficos para visualização de dados
- Cálculos de estatísticas descritivas, como média, mediana, desvio padrão, entre outros
- Análise de regressão e correlação
- Uso de fórmulas e funções para cálculos complexos
- Importação e exportação de dados em diferentes formatos

O Excel é uma ferramenta poderosa para análise de dados, especialmente para aqueles que estão iniciando na área e ainda não estão familiarizados com ferramentas mais avançadas. Com sua interface intuitiva e grande variedade de recursos, o Excel pode ajudar a simplificar a análise de dados e fornecer insights valiosos para tomadas de decisão informadas.



Excel para Análise de Dados - Exemplos



O que é um Banco de Dados Relacional?

*“Um **banco de dados relacional** é um tipo de **banco de dados** que armazena e fornece acesso a pontos de **dados** relacionados entre si. **Bancos de dados relacionais** são baseados no modelo **relacional**, uma maneira intuitiva e direta de representar **dados** em tabelas.” (fonte: [Oracle](#))*



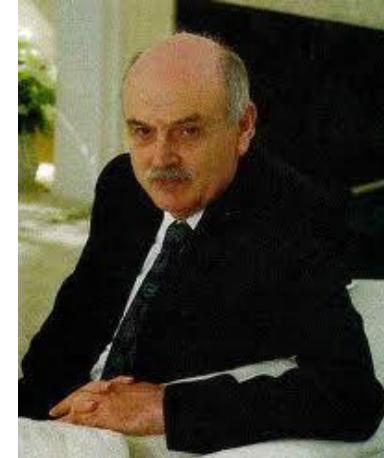
O que são tabela?

*“Tabelas são objetos de banco de dados que contêm todos os dados em um banco de dados. Nas tabelas, os dados são organizados de maneira lógica em um formato de **linha-e-coluna** semelhante ao de uma planilha. Cada linha representa um registro exclusivo e cada coluna representa um campo no registro.” (fonte: [Microsoft](#))*



Onde surgiram os BD? Porque?

- Foi proposto por **Edgar Codd** em 1970
- **Objetivo** de solucionar problemas com armazenamento em arquivos
 - Dificuldade de acesso aos dados
 - Isolamento dos dados
 - Anomalias de acessos concorrentes
 - Problemas com segurança
 - Problemas de integridade
- Vieram os **SGBD's** - Sistemas Gerenciadores de Banco de Dados
 - Armazena os bancos de dados / tabelas...
 - Atende todos estes pontos acima!
 - Segurança | ACID | Acessos
 - Exemplo: SQL Server, PostgreSQL, Oracle...



ACID

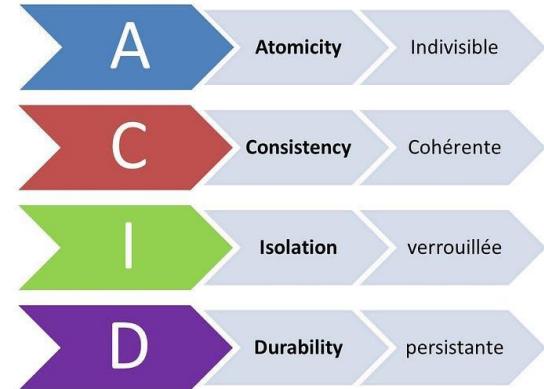
ACID é um conceito que se refere às quatro propriedades de transação de um sistema de banco de dados: **Atomicidade**, **Consistência**, **Isolamento** e **Durabilidade**.

Atomicidade: Em uma transação envolvendo duas ou mais partes de informações discretas, ou a transação será executada totalmente ou não será executada, garantindo assim que as transações sejam atômicas.

Consistência: A transação cria um novo estado válido dos dados ou em caso de falha retorna todos os dados ao seu estado antes que a transação foi iniciada.

Isolamento: Uma transação em andamento mas ainda não validada deve permanecer isolada de qualquer outra operação, ou seja, garantimos que a transação não será interferida por nenhuma outra transação concorrente.

Durabilidade: Dados validados são registrados pelo sistema de tal forma que mesmo no caso de uma falha e/ou reinício do sistema, os dados estão disponíveis em seu estado correto.



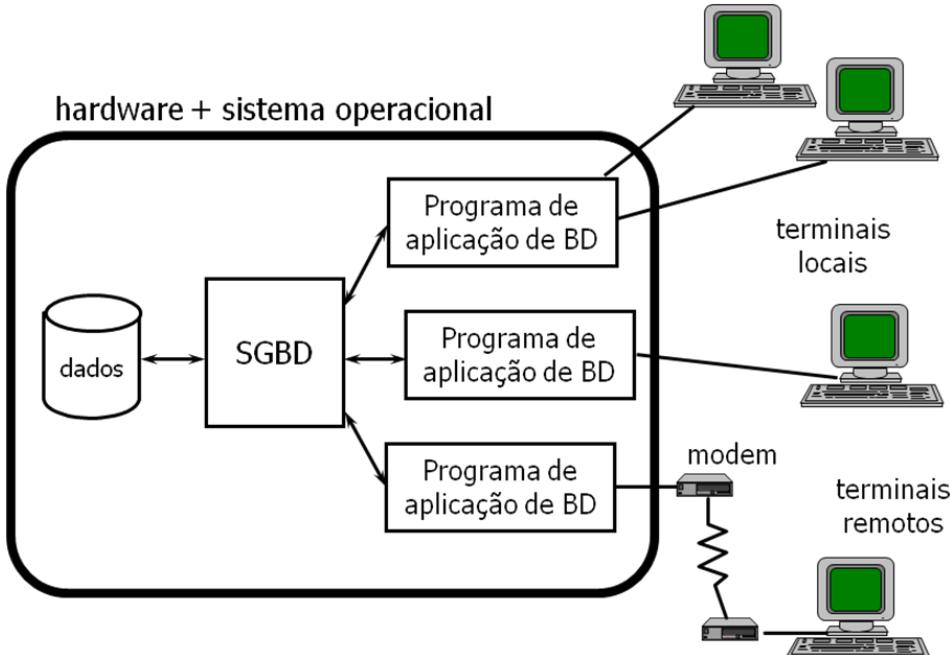
Abstração dos Dados

- **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. (Analistas 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 respeitos. 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. E nem interessa onde estas informações estão armazenadas.



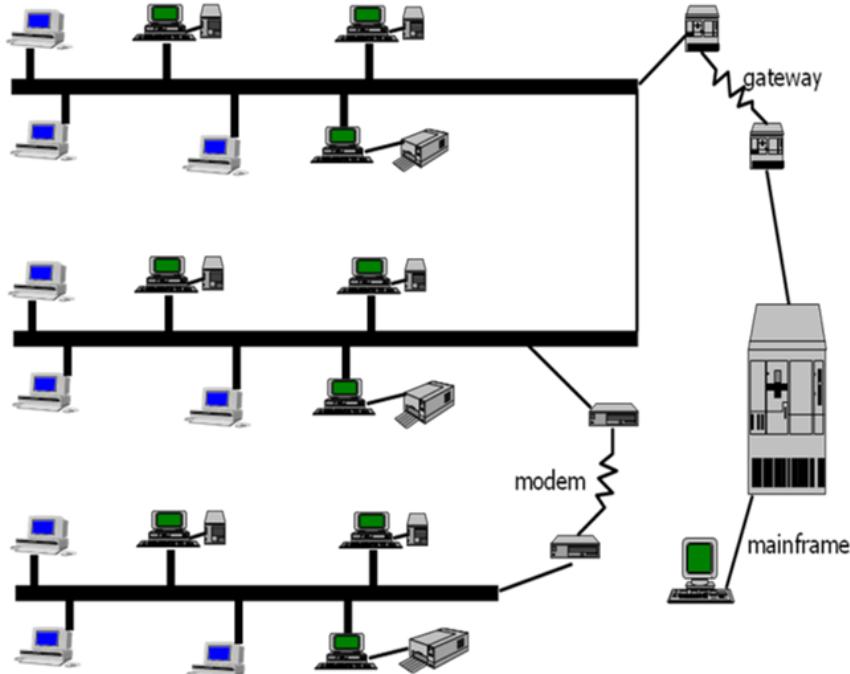
Evolução dos Bancos de Dados

Bancos de Dados Centralizados



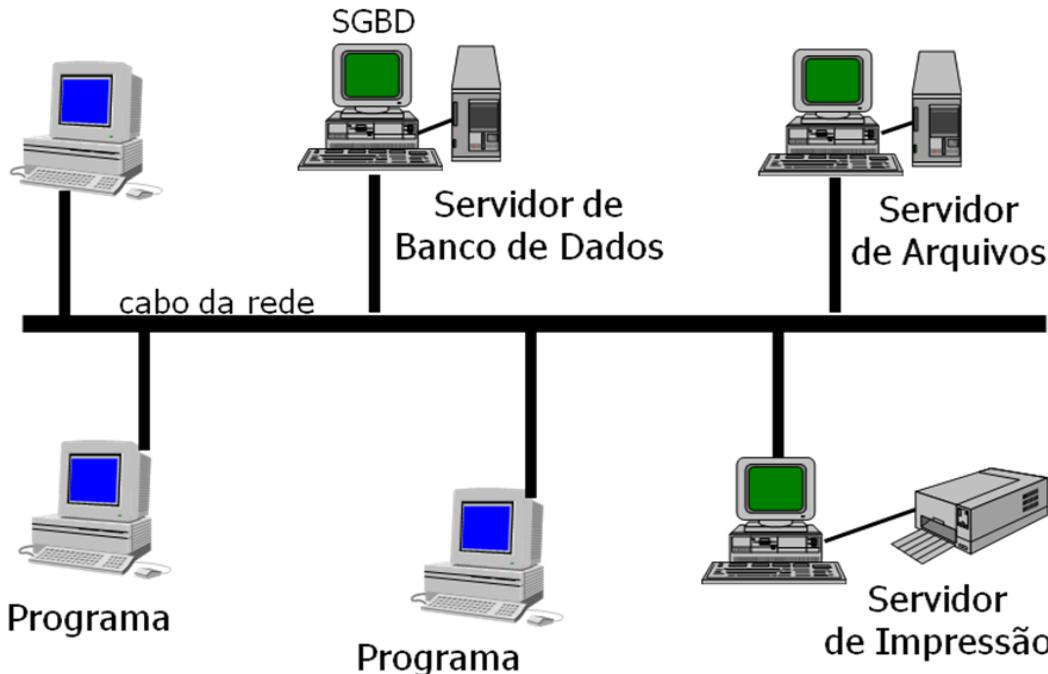
Evolução dos Bancos de Dados

Bancos de Dados Distribuídos



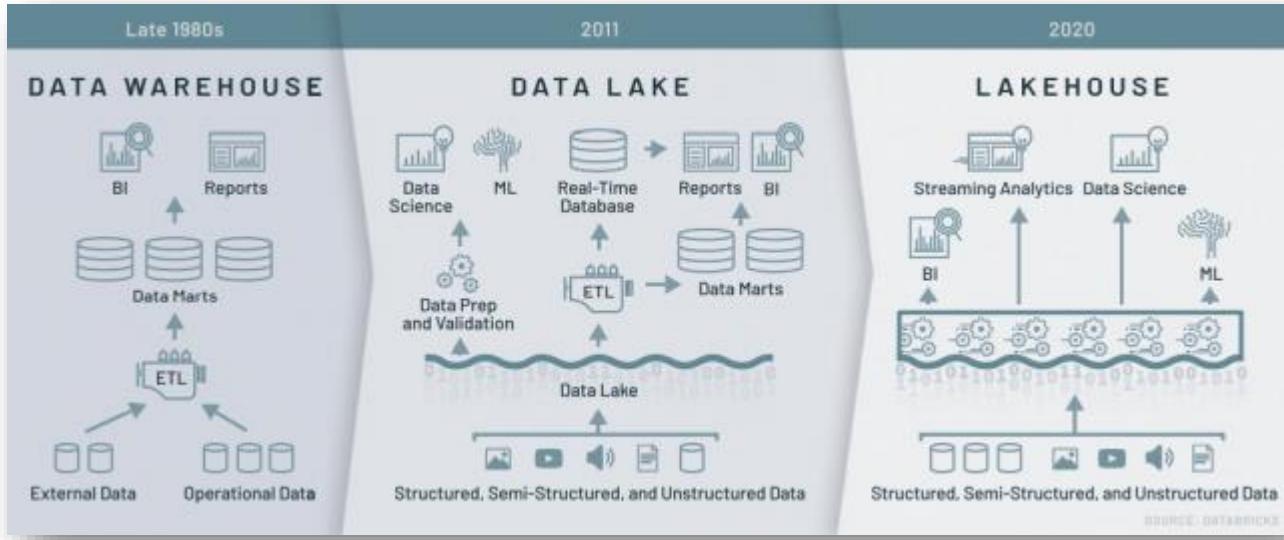
Evolução dos Bancos de Dados

Bancos de Dados Cliente – Servidor (com servidor de Banco de Dados)



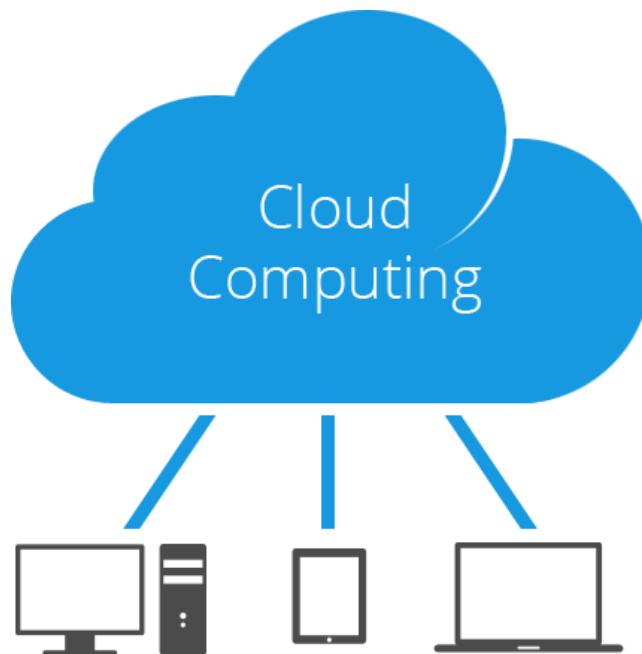
Evolução dos Bancos de Dados

Arquiteturas modernas



Evolução dos Bancos de Dados

Bancos de Dados em nuvem



Amazon Aurora

Vamos montar um banco de dados do zero?

- Levantamento de requisitos (concepção do banco de dados)
- **Modelo Conceitual**
- **Modelo Lógico**
- **Modelo Físico**



Levantamento de Requisitos

Concepção do banco de dados...



Requisitos do cliente



- Preciso de um banco de dados para controlar as vendas diária dos produtos da minha empresa, também precisarei saber se meus vendedores tem dependentes para poder mandar brindes em ocasiões especiais.
- Aqui trabalhamos tanto em loja própria, quanto em loja virtual.



Colhendo requisitos



- Quais informações precisa contar na sua venda?
 - Data
 - Quantidade
 - Valor unitário e total
 - Clientes (nome, sexo, estado, idade)



Colhendo requisitos



- Em relação aos vendedores, Eles recebem comissão?
 - Sim, recebem 
- Em relação aos dependentes, alguma informação específica deseja ter?
 - Sim, o código da Escola que Ele estuda. 
- Qual seria esse código?
 - Código do INEP 



Colhendo requisitos



- Em relação aos produtos que vendem, alguma característica que desejam controlar?
 - Nosso produto tem uma classificação por tipo de produto, tipo A, B... 
- Mais alguma necessidade que deseja nos informar?
 - Normalmente, nossos clientes tem uma classificação com base nas compras que já fizeram, se puderem adicionar... 



Modelo conceitual

“É o modelo de mais alto nível, ou seja, que está mais próximo da realidade dos usuários. O nível conceitual é desenvolvido com alto nível de abstração, a partir dos requisitos do sistema, extraídos na fase de levantamento de requisitos.

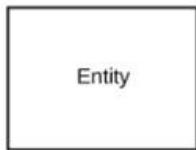
*Esse modelo pode ser elaborado por meio de dois diagramas: **Diagrama de Entidade e Relacionamento (MER)** e/ou o Diagrama de Classes.*



Modelo conceitual

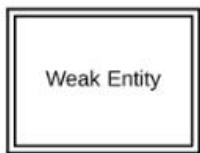
Principais componentes





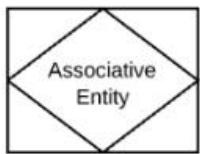
Entity

Entidade forte



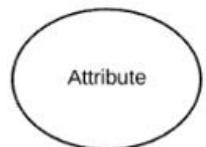
Weak Entity

Entidade fraca



Associative Entity

Entidade associativa



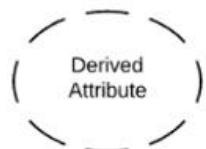
Attribute

Atributo



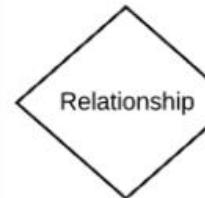
Multivalued Attribute

Atributo multivvalorado



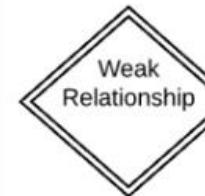
Derived Attribute

Atributo derivado



Relationship

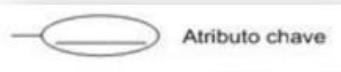
Relacionamento



Weak Relationship

Relacionamento fraco

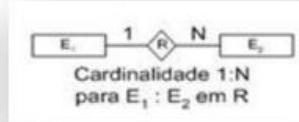
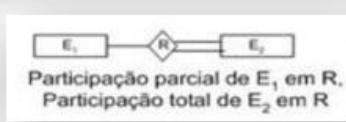
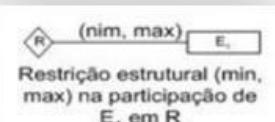
Conectividade	Peter Chan	James Martin
1:1		
1:N		
N:N		



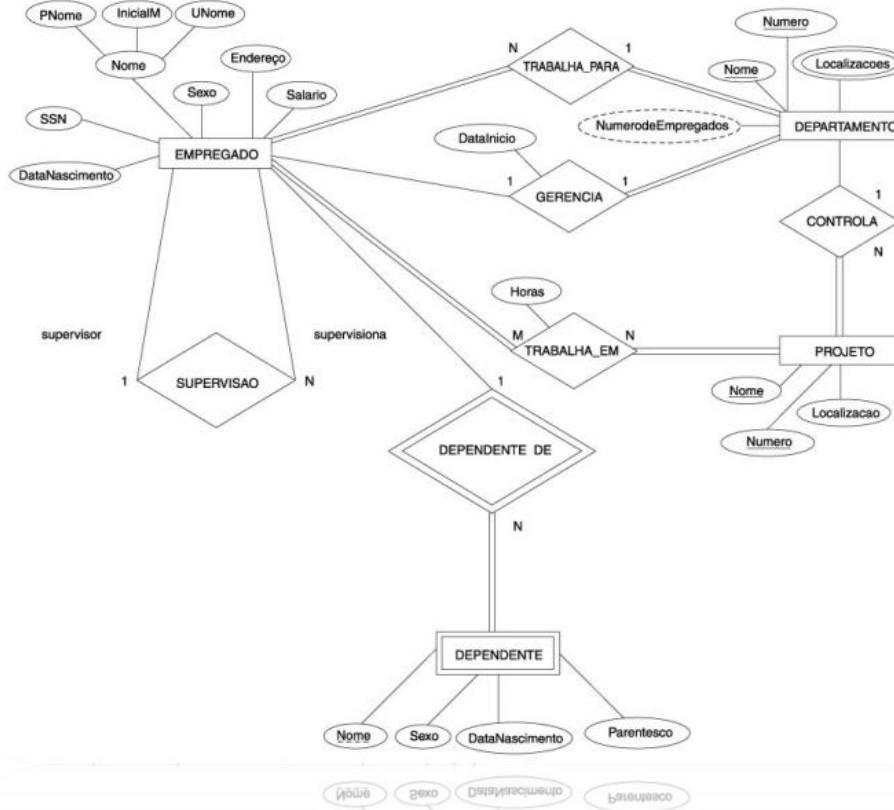
Atributo chave



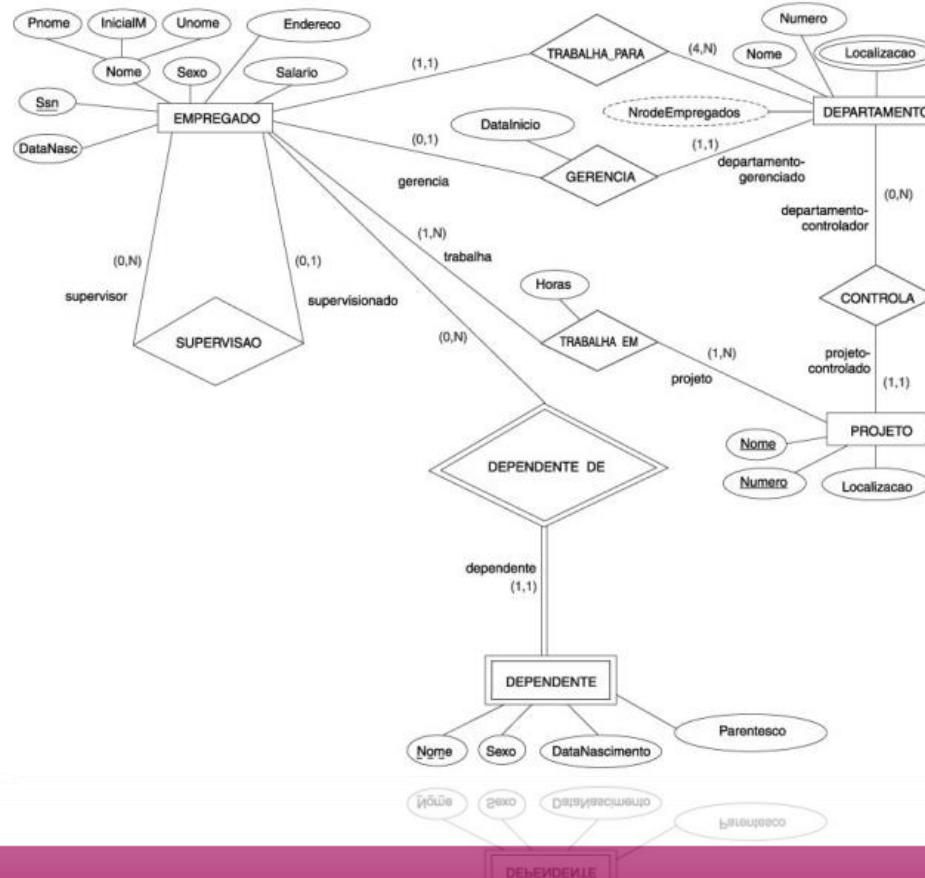
Atributo composto

Cardinalidade 1:N
para E₁: E₂ em RParticipação parcial de E₁ em R.
Participação total de E₂ em RRestrição estrutural (min, max)
na participação de E₁ em R

Exemplo 1 - MER

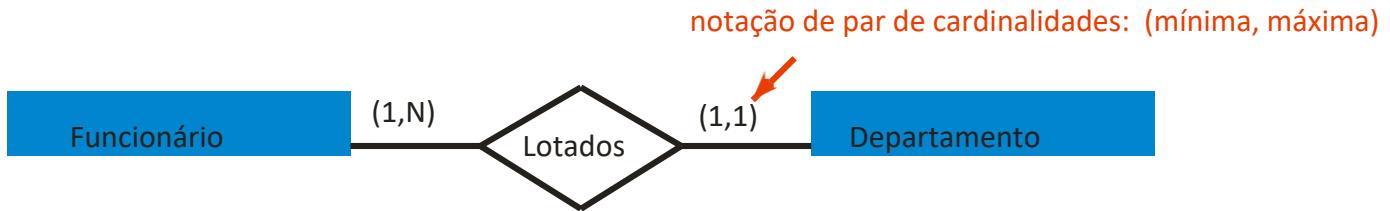


Exemplo 2 – MER com Notação (min,max)



Cardinalidade com Notação (min,max)

Indica se a participação das ocorrências de entidades no relacionamento é obrigatória ou opcional



Um funcionário pode estar lotado em 1 departamento e exclusivamente em 1.

Um departamento obrigatoriamente tem que ter de 1 até N empregados lotados nele

Modelo de Entidade e Relacionamento (MER)

- Pelo que coletamos do cliente, no nosso banco de dados, teremos 4 entidades
 - Vendas
 - Vendedores
 - Dependentes
 - Produto



Modelo de Entidade e Relacionamento (MER)

- Vendas (tbven)
 - cdven (pk)
 - dtven
 - cdcli
 - nmcli
 - agecli
 - clacli
 - sxcli
 - cidcli
 - estcli
 - paiscli
 - qtven
 - vruven
 - vrven
 - canal
 - stven
 - deleted
 - cdvdd* (fk)
 - cdpro* (fk)

Requisitos do cliente

- Preciso de um banco de dados para controlar as vendas diárias dos produtos da minha empresa, também preciso saber se meus vendedores tem dependentes para poder mandar brindes em ocasiões especiais.
- Aqui trabalhamos tanto em loja própria, quanto em loja virtual.

Colhendo requisitos

- Quais informações precisa contar na sua venda?
 - Data
 - Quantidade
 - Valor unitário e total
 - Clientes (nome, sexo, estado, idade)



Modelo de Entidade e Relacionamento (MER)

- Vendedores (tbvdd)
 - cdvdd (pk)
 - nmvdd
 - sxvdd
 - perccomissao
 - matfunc

Requisitos do cliente

- Preciso de um banco de dados para controlar as vendas diárias dos produtos da minha empresa, também preciso saber se meus vendedores tem dependentes para poder mandar brindes em ocasiões especiais.
- Aqui trabalhamos tanto em loja própria, quanto em loja virtual.

Colhendo requisitos

- Em relação aos vendedores, Eles recebem comissão?
 - Sim, recebem 🧑
- Em relação aos dependentes, alguma informação específica deseja ter?
 - Sim, o código da Escola que Ele estuda. 🧑
 - Qual seria esse código?
 - Código do INEP 🧑

Modelo de Entidade e Relacionamento (MER)

- Dependentes (tbdep)
 - cddep (pk)
 - nmdep
 - dtnasc
 - sxdep
 - inepescola
 - cdvdd (fk)

Requisitos do cliente

- Preciso de um banco de dados para controlar as vendas diárias dos produtos da minha empresa, também preciso saber se meus vendedores tem dependentes para poder mandar brindes em ocasiões especiais.
- Aqui trabalhamos tanto em loja própria, quanto em loja virtual.

Colhendo requisitos

- Em relação aos vendedores, Eles recebem comissão?
 - Sim, recebem 🚧
- Em relação aos dependentes, alguma informação específica deseja ter?
 - Sim, o código da Escola que Ele estuda. 🚧
 - Qual seria esse código?
 - Código do INEP 🚧

Modelo de Entidade e Relacionamento (MER)

- Produto (tbpro)
 - cdpro (pk)
 - nmpro
 - tppro
 - undpro
 - slpro
 - stpro

Requisitos do cliente

- Preciso de um banco de dados para controlar as vendas diárias dos produtos da minha empresa, também preciso saber se meus vendedores tem dependentes para poder mandar brindes em ocasiões especiais.
- Aqui trabalhamos tanto em loja própria, quanto em loja virtual.

Colhendo requisitos

- Em relação aos produtos que vendem, alguma característica que desejam controlar?
 - Nosso produto tem uma classificação por tipo de produto, tipo A, B,.. 
- Mais alguma necessidade que deseja nos informar?
 - Normalmente, nossos clientes tem uma classificação com base nas compras que já fizeram, se puderem adicionar... 

BR Modelo Web

Português do Brasil | English

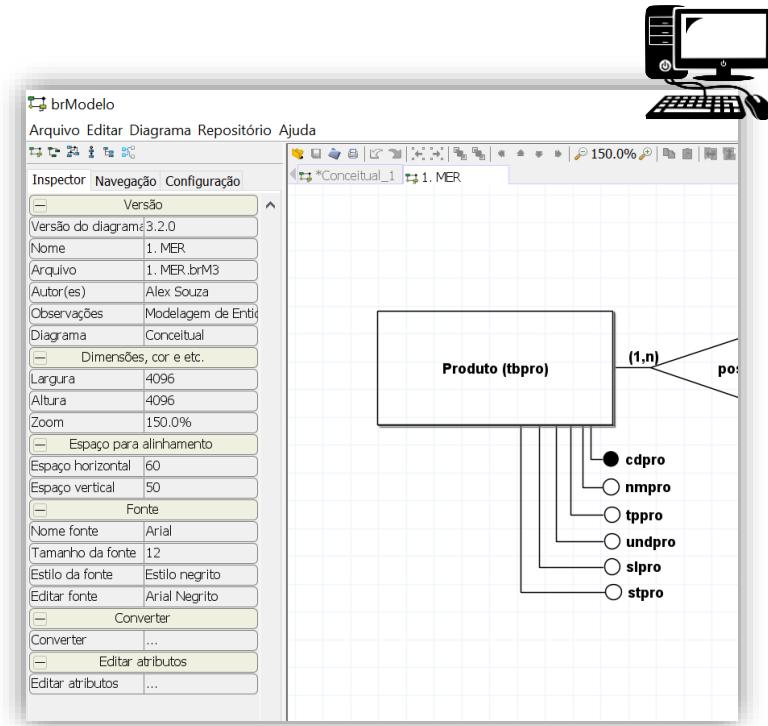
Ferramenta para modelagem de banco de dados

100% online e gratuita

Acesse Crie sua conta

Modelagem conceitual de: Recursos Humanos

<https://app.brmodeloweb.com/>

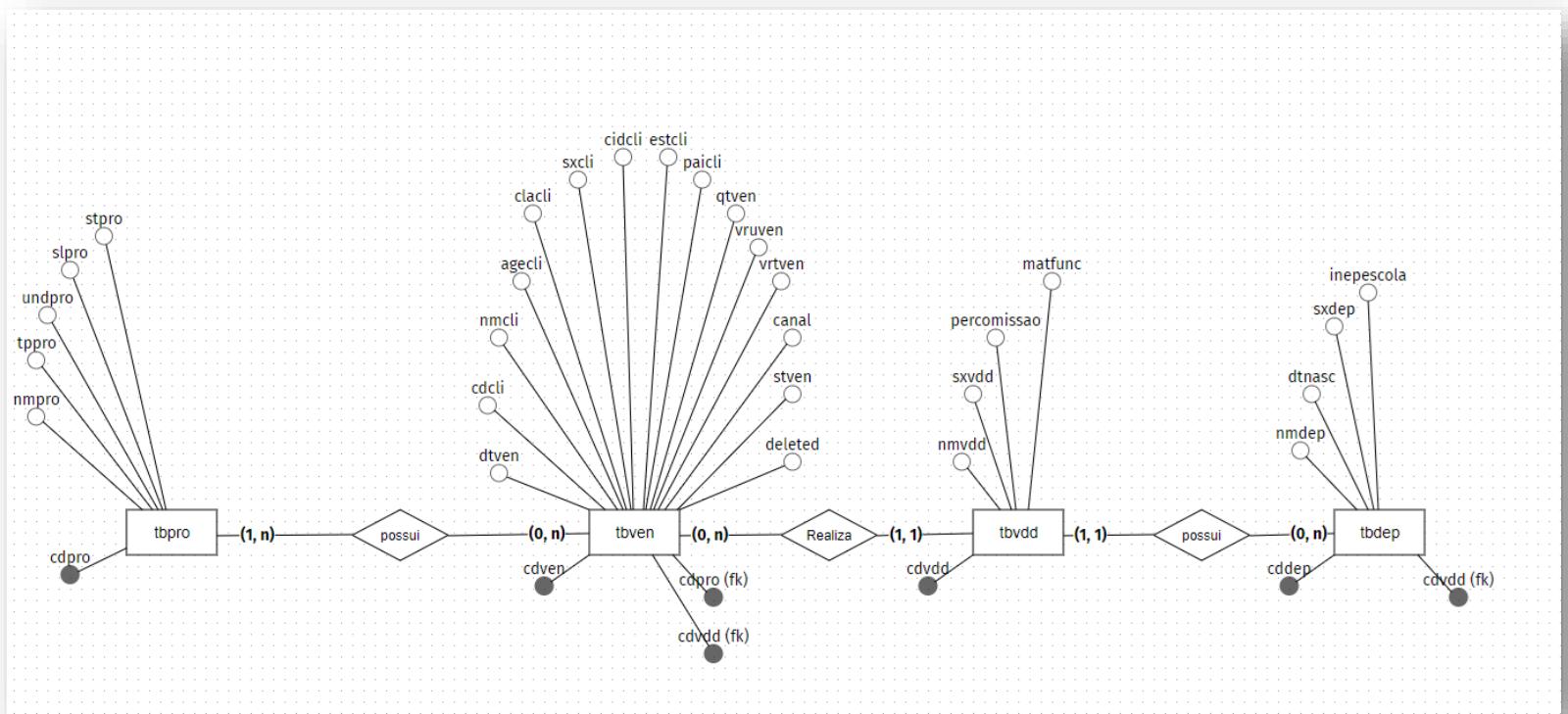


Desafio

Criar a Modelagem Conceitual do nosso banco de dados



Modelo de Entidade e Relacionamento (MER)



Alinhamentos (20.03.2023)



- Discord (uso)
- Formulário de alinhamento
- GitHub
 - Criação da conta
 - Formulário de alinhamento
 - Alguém com dificuldade aqui?
- Networking entre vocês
- “Monitores da turma”
- Escuta Ativa
- Atividades
 - Modelo Conceitual
 - Modelo Lógico
 - Maturidade de Dados
 - Leitura e formulário
- (agora) Troca de ideias sobre Modelagem

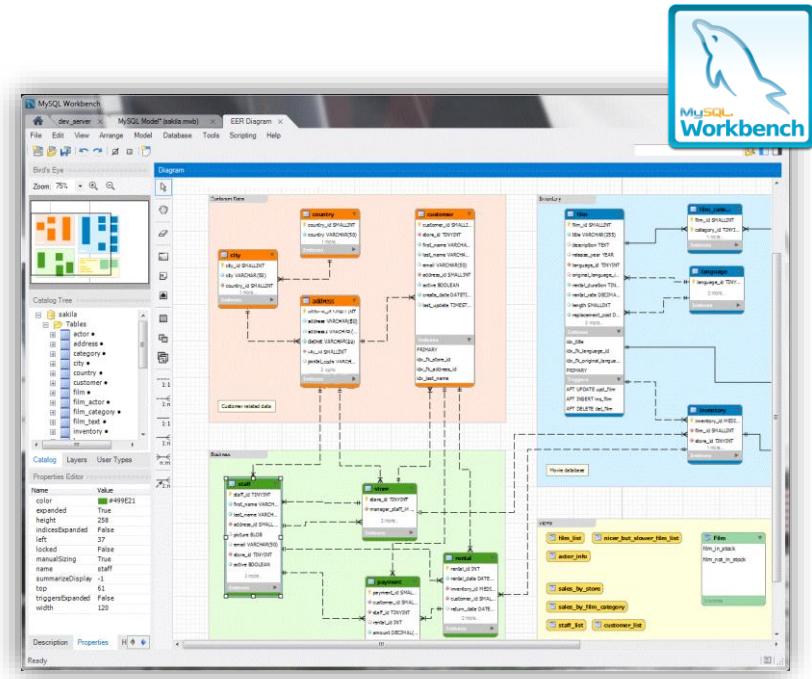


Modelo lógico

“Descreve como os dados serão armazenados no banco e também seus relacionamentos.

*Esse modelo adota alguma tecnologia, pode ser: relacional, orientado a objetos,
orientado a colunas, entre outros.*

Já é modelo que ficará o banco de dados, informando os atributos, tipos de dados...



<https://app.brmodeloweb.com/>

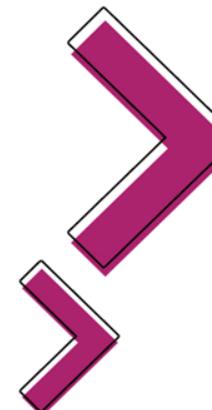
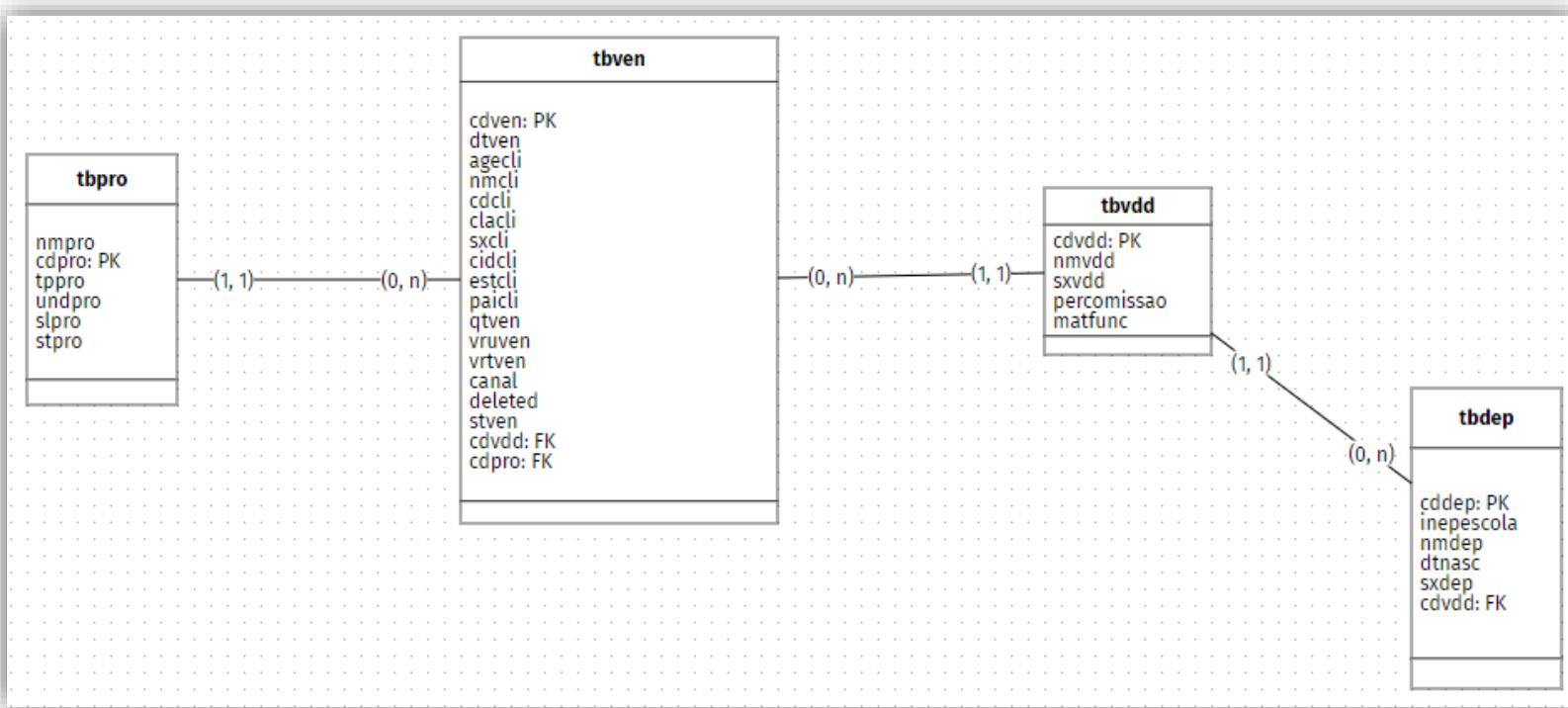


Desafio

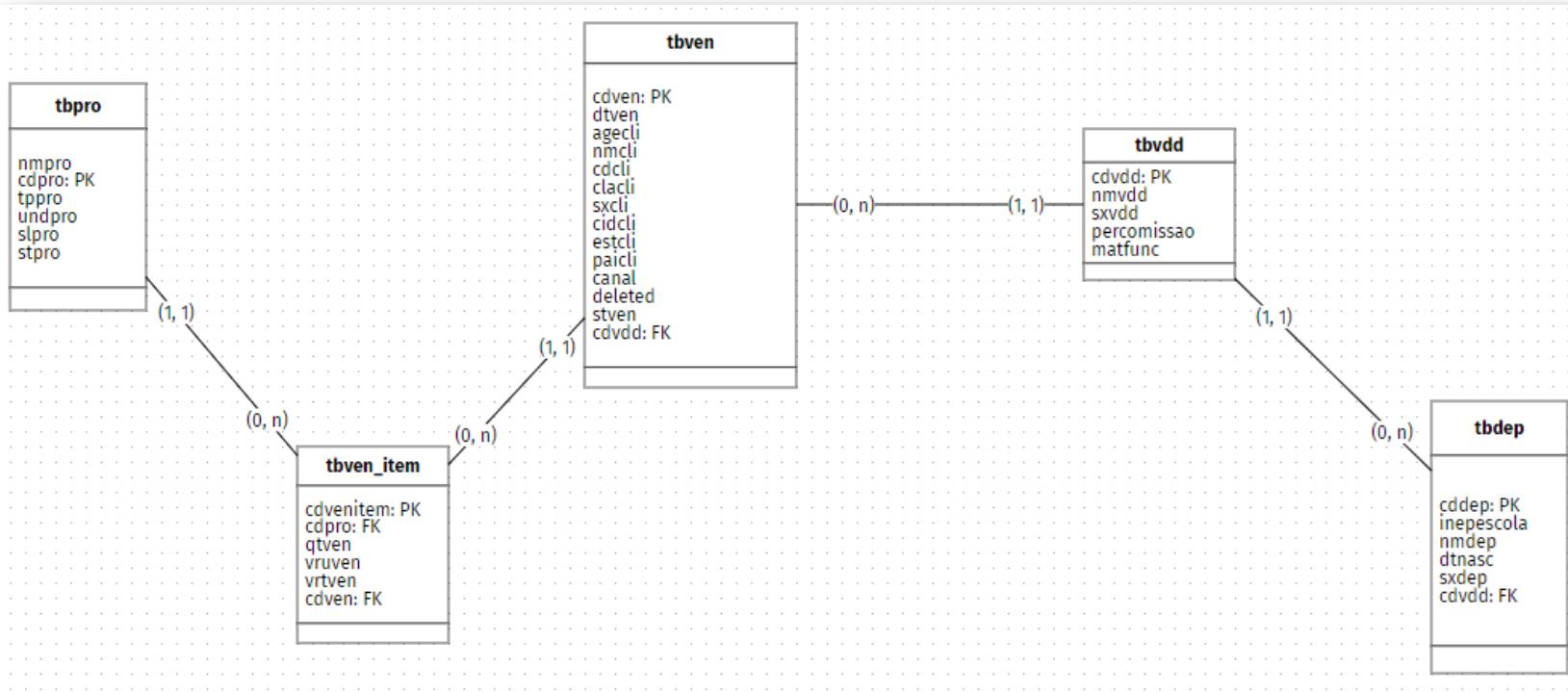
Criar a Modelagem Lógica do nosso banco de dados



Modelo Lógico

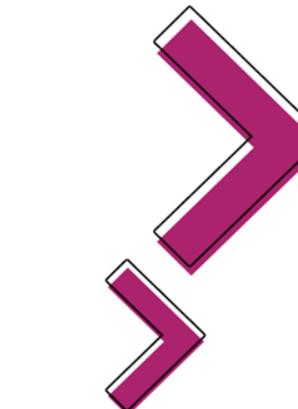


Modelo Lógico (atendendo ao conceitual)



Modelo Físico

“Descreve, por meio de alguma linguagem, como será feita a armazenagem no banco. Nesse nível se escolhe qual Sistema gerenciador de Banco de dados (SGBD) será usado, levando em consideração o modelo lógico adotado. Pode ser: SQL Server, PostgreSQL, MySQL, dentre outros.”*



Revisão



Quais as 4 etapas para criação de um Banco de Dados?

Requisitos, Modelagem Conceitual,
Modelagem Lógica e Modelagem
Física



O que é feito na fase de Requisitos?



O que é feito na Modelagem Lógica?

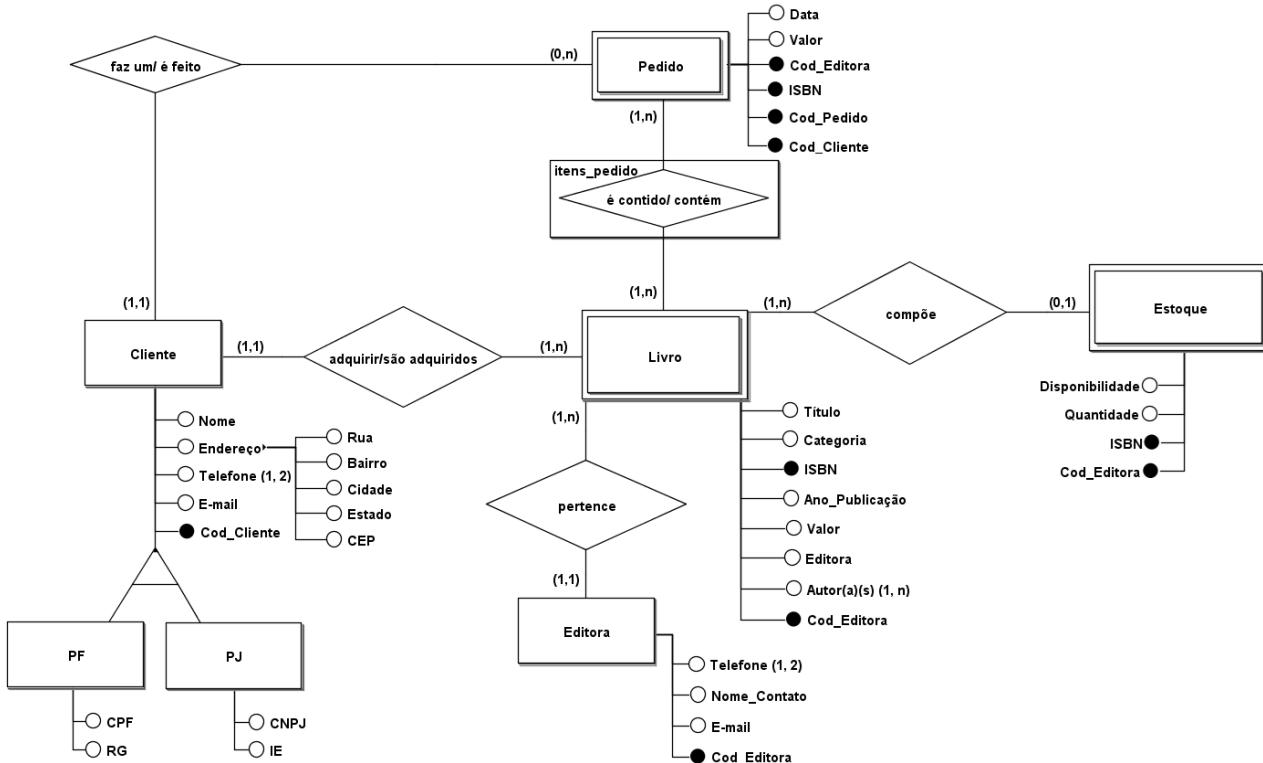


O que é feito na Modelagem Conceitual?



**O que é um SGBD? Qual iremos
utilizar aqui? Me dê alguns outros
exemplos...**

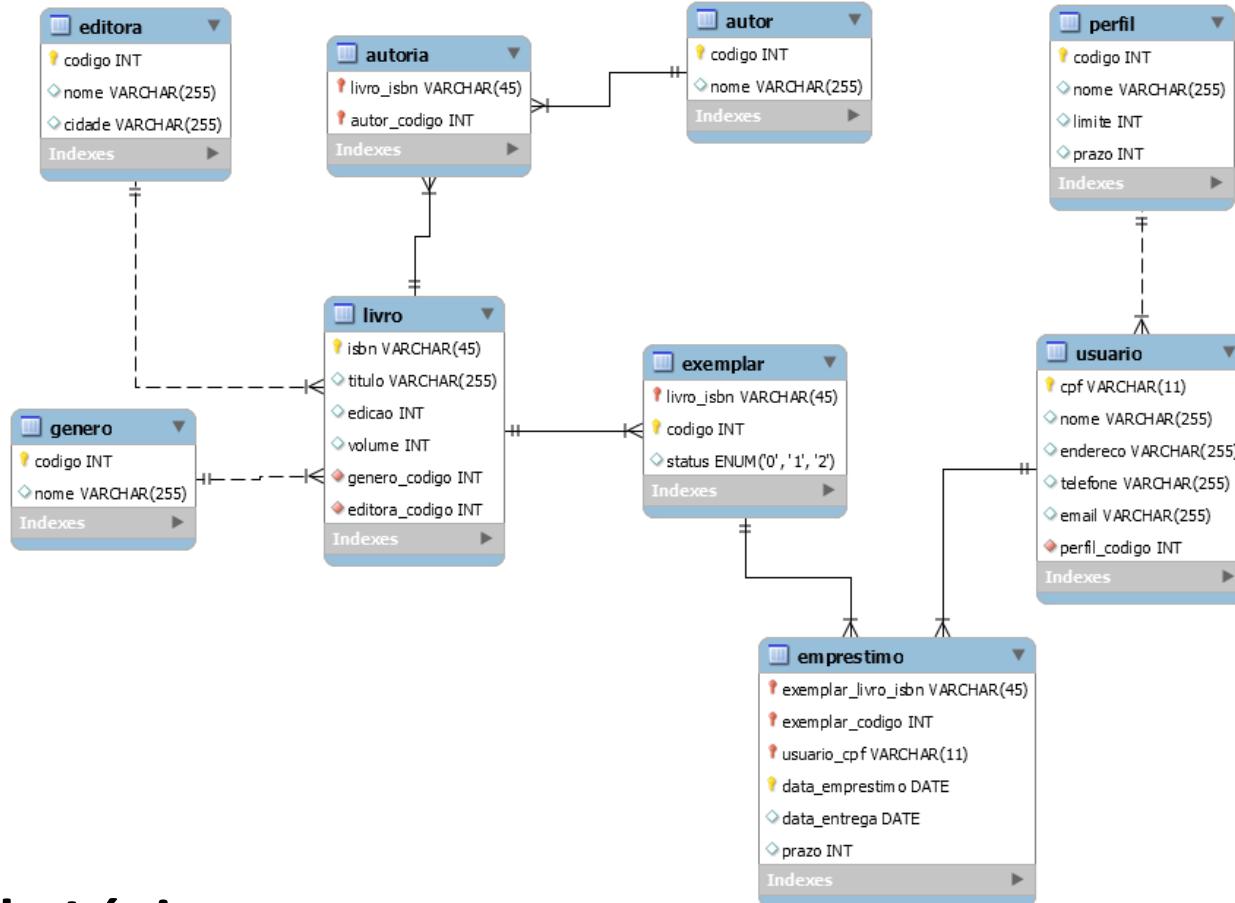




Modelo Conceitual

Porque habilidades comportamentais é importante?





Modelo Lógico

O que é uma tabela?



**Como as tabelas podem se
relacionar?**



Vamos por camadas...

SGBD Segurança, concorrência, ACID...

BANCOS DE DADOS

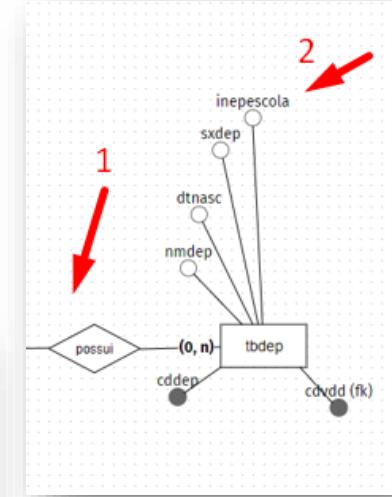
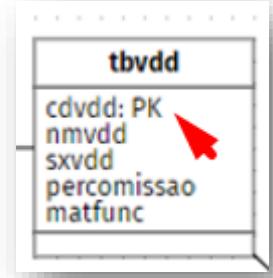
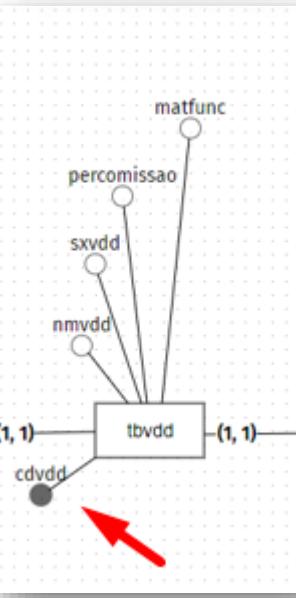
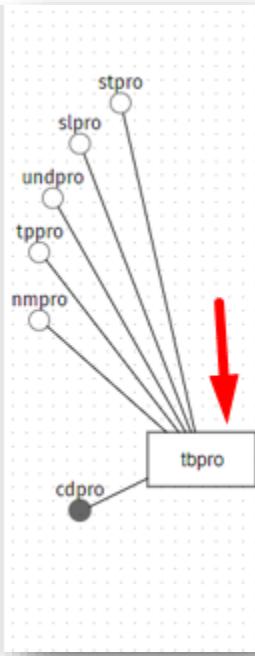
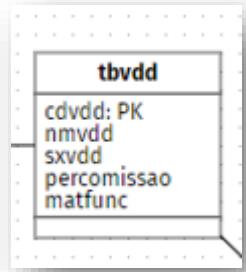
TABELAS

LINHAS e COLUNAS

TIPOS DE DADOS

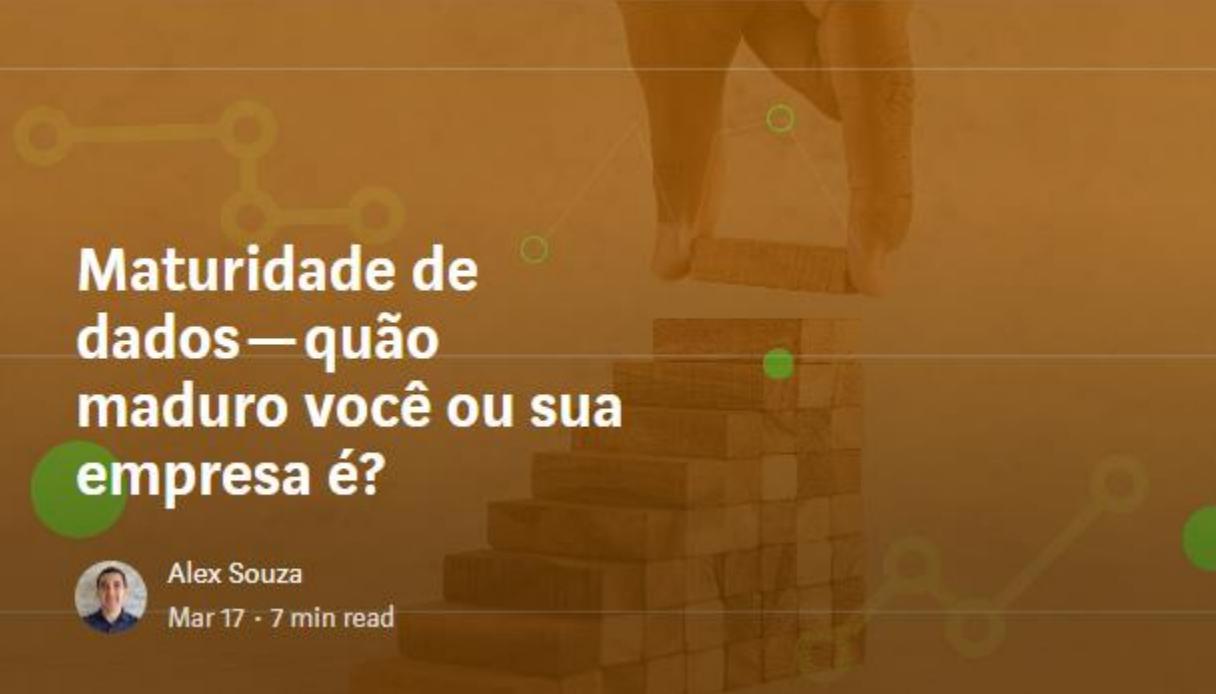


O que é isso?



Versionamento
Documentação
Portfólio de Serviços





Maturidade de dados—quão maduro você ou sua empresa é?



Alex Souza

Mar 17 · 7 min read

[Link para o formulário](#)



Leitura da apostila



O que estão achando até o
momento?



O que é feito na Modelagem Física?



Fim Revisão



Visão Geral – Banco de Dados

Documentação do PostgreSQL



PostgreSQL



Em
Breve!



COMUNIDADE DE
ANÁLISE DE DADOS

MERCADO e SKILLS

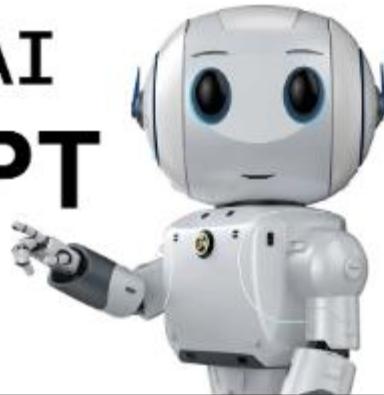
Quinta-feira (30.03) - 19h - Zoom



ELIESER LOPES



OpenAI
ChatGPT



Tipos da Linguagem SQL

- **DDL - Data Definition Language - Linguagem de Definição de Dados.**
 - São os comandos que interagem com os objetos do banco. (**CREATE, ALTER e DROP**)
- **DML - Data Manipulation Language - Linguagem de Manipulação de Dados.**
 - São os comandos que interagem com os dados dentro das tabelas. (**INSERT, DELETE e UPDATE**)
- **DTL - Data Transaction Language - Linguagem de Transação de Dados.**
 - São os comandos para controle de transação. (**BEGIN TRANSACTION, COMMIT E ROLLBACK**)
- **DCL - Data Control Language - Linguagem de Controle de Dados.**
 - São os comandos para controlar a parte de segurança do banco de dados. (**GRANT, REVOKE E DENY**)
- **DQL - Data Query Language - Linguagem de Consulta de dados.**
 - São os comandos de consulta. (**SELECT**)
 - *Em alguns livros o **SELECT** fica na **DML** em outros tem esse grupo próprio.



Modelo Físico

*Criando o banco de dados e as tabelas com base no
Modelo Lógico e Populando as tabelas*

Nome do BD: ***DC_DA_1_SeuNome***



PostgreSQL



aasouzaconsult / DC_Data-Analytics_01 Private

Code Issues Pull requests Actions Projects Security Insights Settings

main · 1 branch · 0 tags

Go to file Add file <> Code

aasouzaconsult Update readme.md b854dd4 11 minutes ago 50 commits

Apostilas Delete APOSTILA DATA ANALYTICS U2 M1 - PROCESSO DE DESCOBERT... 3 weeks ago

Arquivos Add files via upload 3 weeks ago

Excel para Análise de Dados Update README.md 3 weeks ago

Instalação PostgreSQL 14 em Imagens Update README.md 3 weeks ago

Scripts SQL Update readme.md 11 minutes ago

Slides Aula Add files via upload 4 days ago

README.md Update README.md yesterday

README.md

Formação de Data Analytics - Turma 1 (Sul) - Digital College



PostgreSQL



É
dia 6!



COMUNIDADE DE
ANÁLISE DE DADOS

Habilidades Comportamentais e Propósito



LÍDIA ARÁUJO

Quinta-feira (06.04) - 19h - Zoom

Componentes do Modelo Relacional

Entidade = 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

Componentes do Modelo Relacional

Atributo = Coluna

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

Atributos

Componentes 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

Componentes 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



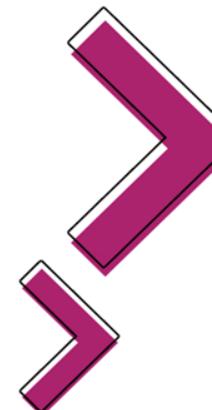
Componentes do Modelo Relacional

Tupla = Linha

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



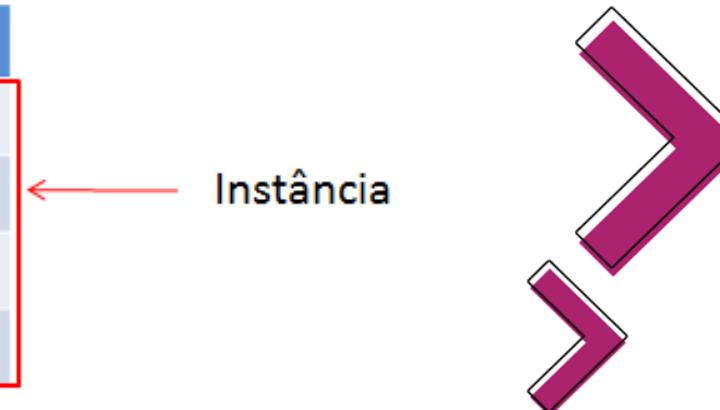
Tupla



Componentes 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



Componentes do Modelo Relacional

Campo = Encontro de Linha e Coluna

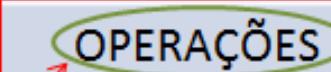
TABELA DEPARTAMENTO		
ID	NOME	ID_REGIAO
10	VENDAS	1
20	OPERAÇÕES	1
30	ADMINISTRATIVO	2
40	FINANCEIRO	3

Campo



Componentes 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

Componentes 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

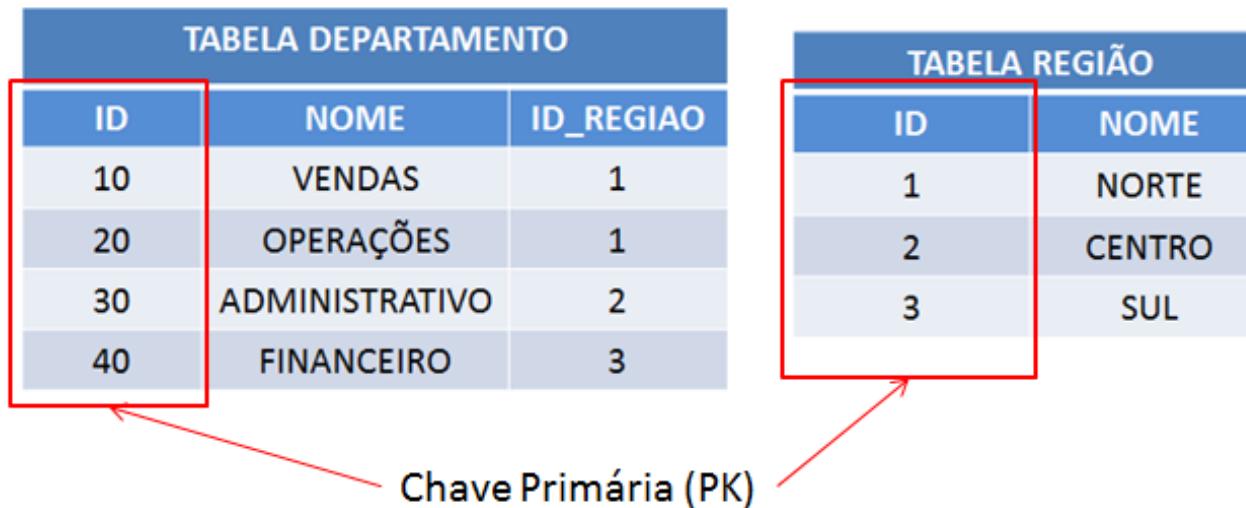
Componentes do Modelo Relacional

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)



Componentes do Modelo Relacional

Chave Primaria = 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

Componentes do Modelo Relacional

Chave Primaria = Atributo Único

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

A coluna "ID" é circundada por um retângulo vermelho e contém os valores 10, 20, 20 e 40. O valor 20 aparece duas vezes e é circulado por um círculo vermelho. Um ponto de seta aponta da palavra "PK" para o lado esquerdo da coluna "ID".

Integridade de Chave Primária

Componentes do Modelo Relacional

Chave Primaria = Atributo Obrigatório

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 Entidade

Componentes do Modelo Relacional

Chave Primaria = Atributo Obrigatório

TABELA REGIÃO	
ID	NOME
	NORTE
2	CENTRO
	SUL

Nulo

PK

Integridade de Entidade

Componentes do Modelo Relacional

Chave Primaria Simples

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)

Componentes do Modelo Relacional

Chave Primaria 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

Componentes do Modelo Relacional

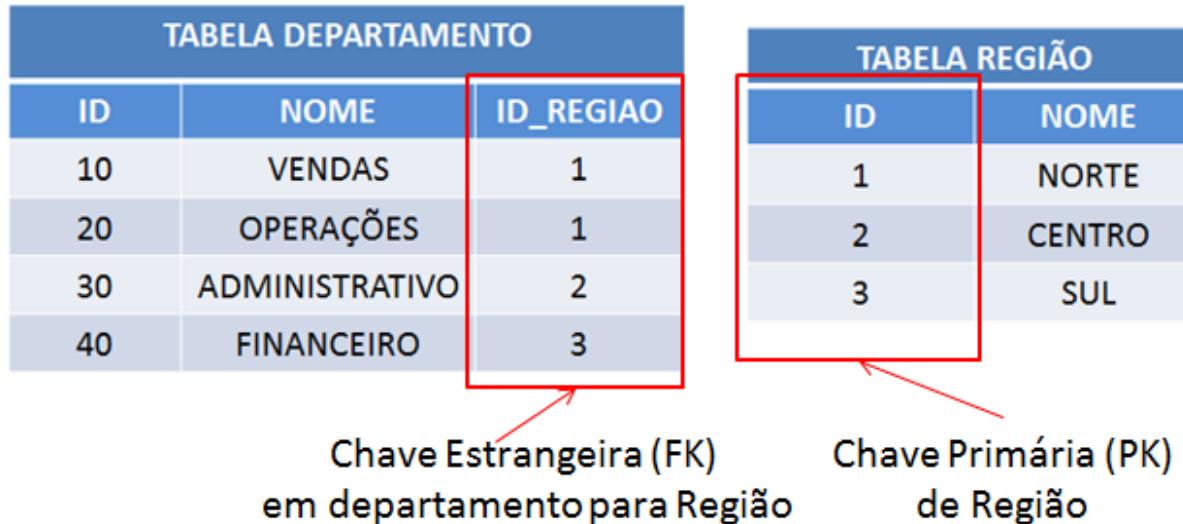
Chave Candidata

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

Chaves Candidatas

Componentes do Modelo Relacional

Chave Estrangeira



Chave Estrangeira -> Integridade Referencial

Modelo Físico

Consultando as tabelas e Relacionando tabelas



Dicas de boas práticas

- *Comentários*
- *Indentação*
- *Escrita Limpa*
- *Padronização de objetos*
- *Evite * (select)*



Tipos da Linguagem SQL

- **DDL - Data Definition Language** - Linguagem de Definição de Dados.
 - São os comandos que interagem com os objetos do banco. (CREATE, ALTER e DROP)
- **DML - Data Manipulation Language** - Linguagem de Manipulação de Dados.
 - São os comandos que interagem com os dados dentro das tabelas. (INSERT, DELETE e UPDATE)
- **DTL - Data Transaction Language** - Linguagem de Transação de Dados.
 - São os comandos para controle de transação. (BEGIN TRANSACTION, COMMIT E ROLLBACK)
- **DCL - Data Control Language** - Linguagem de Controle de Dados.
 - São os comandos para controlar a parte de segurança do banco de dados. (GRANT, REVOKE E DENY)
- **DQL - Data Query Language** - Linguagem de Consulta de dados.
 - São os comandos de consulta. (SELECT)
 - *Em alguns livros o SELECT fica na DML em outros tem esse grupo próprio.

SQL

Conhecendo principais comandos...



DDL – Data Definition Language

- CREATE
- ALTER
- DROP

> -

DML – Data Manipulation Language

- INSERT
- UPDATE
- DELETE



DQL – Data Query Language

- SELECT
 - FROM
 - WHERE



DQL - Operadores

- *
- DISTINCT
- AS (ALIAS)
- LIMIT
- NULL
- BETWEEN
- ORDER BY
- HAVING
- CAST
- CONVERT



DQL – Operadores+

- UNION
- INTERSECT
- EXCEPT
- DATE

> -

DQL – Operadores+

- LIKE
- UPPER
- LOWER
- TRIM
- REPLACE
- LPAD (REPLICATE)
- RPAD
- ...



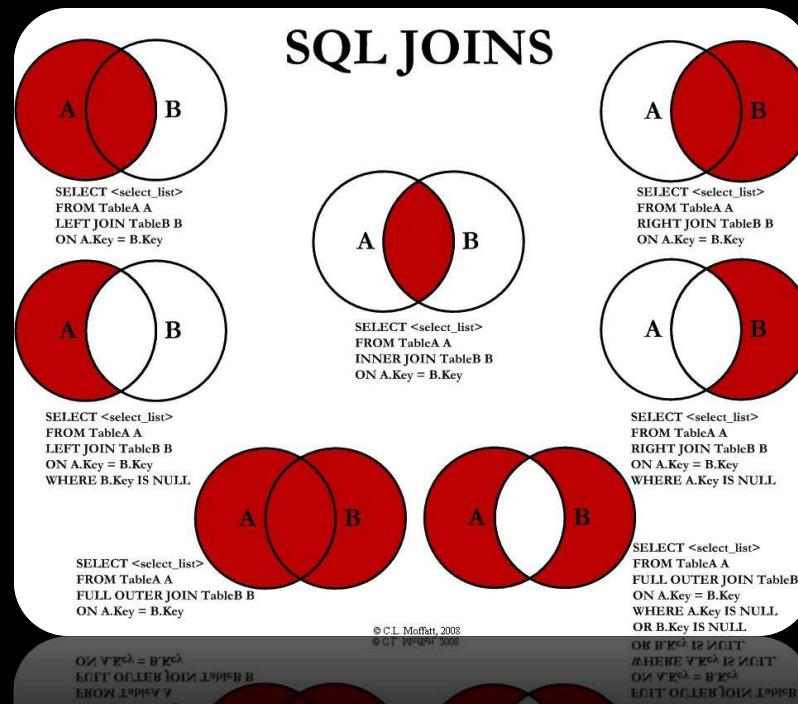
DQL – Operadores Agregação

- SUM *GROUP BY
- AVG
- MAX
- MIN
- COUNT



DQL - Relacionamentos (JOINS)

- INNER JOIN
- LEFT JOIN
- RIGHT JOIN
- OUTER JOIN

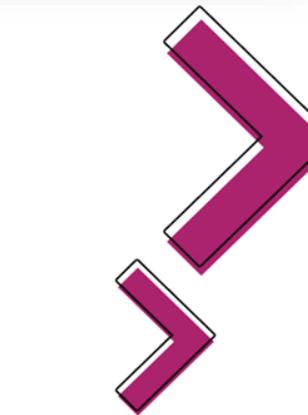


Tipos da Linguagem SQL

- **DDL - Data Definition Language** - Linguagem de Definição de Dados.
 - São os comandos que interagem com os objetos do banco. (CREATE, ALTER e DROP)
- **DML - Data Manipulation Language** - Linguagem de Manipulação de Dados.
 - São os comandos que interagem com os dados dentro das tabelas. (INSERT, DELETE e UPDATE)
- **DTL - Data Transaction Language** - Linguagem de Transação de Dados.
 - São os comandos para controle de transação. (BEGIN TRANSACTION, COMMIT E ROLLBACK)
- **DCL - Data Control Language** - Linguagem de Controle de Dados.
 - São os comandos para controlar a parte de segurança do banco de dados. (GRANT, REVOKE E DENY)
- **DQL - Data Query Language** - Linguagem de Consulta de dados.
 - São os comandos de consulta. (SELECT)
 - *Em alguns livros o SELECT fica na DML em outros tem esse grupo próprio.

SQL

Praticando principais comandos...

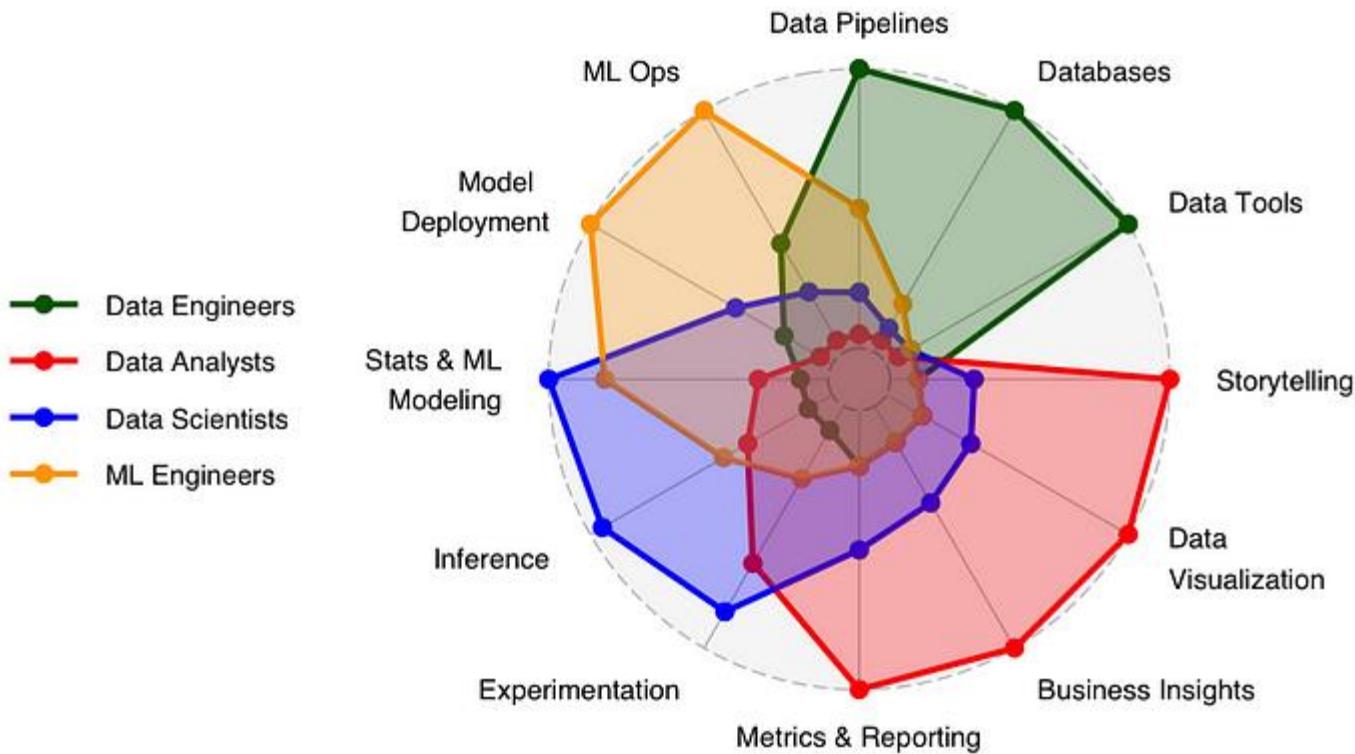


Desafio

Criar a documentação do Banco de Dados

Deverá conter:

- *O que foi especificado pelo cliente*
- *O desenho da modelagem conceitual e lógica*
- *A descrição de cada coluna e seu tipo de dados*
- *Algumas consultas modelo*





SQL

Normalização de Dados (Formas Normais)



Desafio

Normalização de Dados (Formas Normais)

Normalizar a tabela: tbven

Entregar (github): script e documentação informando qual ou quais formas normais utilizou.



Tipos da Linguagem SQL

- DDL - *Data Definition Language* - Linguagem de Definição de Dados.
 - São os comandos que interagem com os objetos do banco. (CREATE, ALTER e DROP)
- DML - *Data Manipulation Language* - Linguagem de Manipulação de Dados.
 - São os comandos que interagem com os dados dentro das tabelas. (INSERT, DELETE e UPDATE)
- DTL - *Data Transaction Language* - Linguagem de Transação de Dados.
 - São os comandos para controle de transação. (BEGIN TRANSACTION, COMMIT E ROLLBACK)
- DCL - *Data Control Language* - Linguagem de Controle de Dados.
 - São os comandos para controlar a parte de segurança do banco de dados. (GRANT, REVOKE E DENY)
- DQL - *Data Query Language* - Linguagem de Consulta de dados.
 - São os comandos de consulta. (SELECT)
 - *Em alguns livros o SELECT fica na DML em outros tem esse grupo próprio.

SQL

Praticando um pouco mais de DDL e DML

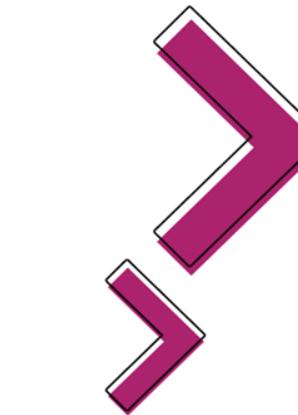
Passos:

- **Criar uma tabela: tbpessoa com os campos: cod, nome, cpf e datanasc**
- **Inserir 10 registros aleatórios (dados fictícios)**
- **Alterar a datanasc do cod de número 5 para 08/03/2022**
- **Deletar o registro do cod de número 10**
- **Alterar o tipo de dado do campo: cod para smallint, se tiver smallint, mude para bigint**
- ***** atenção ao Where *****



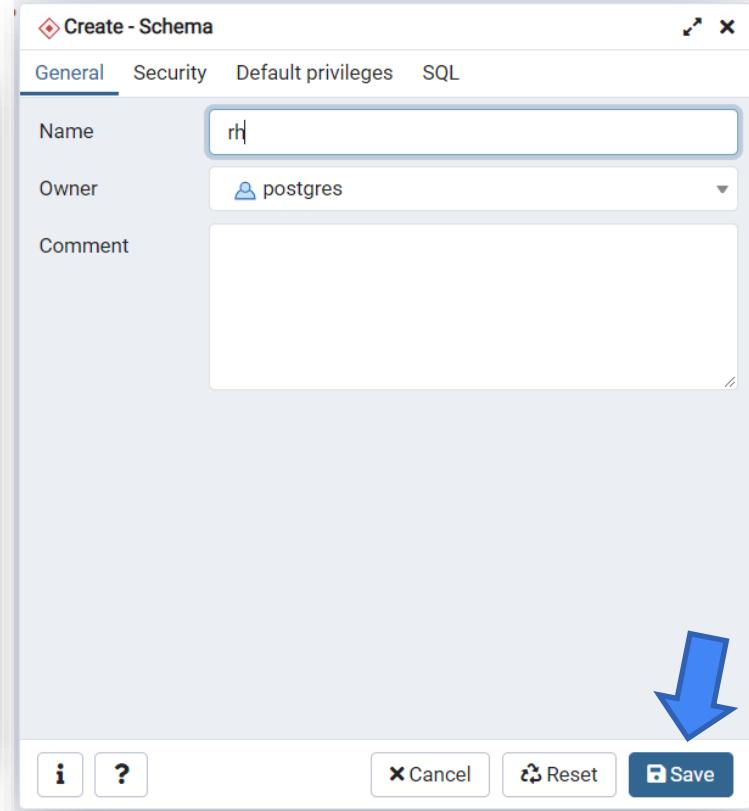
PostgreSQL

Diversas

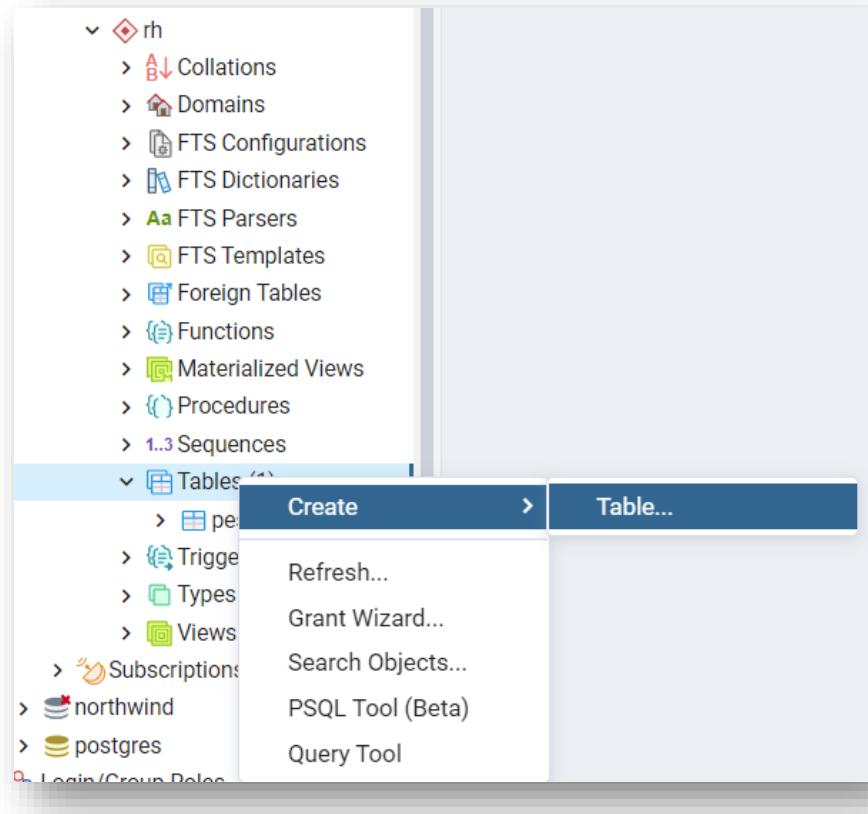


Criando Schemas

The screenshot shows the pgAdmin interface for PostgreSQL 14. On the left, there's a tree view of database objects under 'PostgreSQL 14'. A right-click context menu is open over the 'Schemas (1)' node. The menu items are: 'Create', 'Schema...', 'Refresh...', 'Search Objects...', 'PSQL Tool (Beta)', and 'Query Tool'. The 'Create' item is highlighted.



Criando tabelas (forma gráfica)



Adicionando colunas (forma gráfica)

Create - Table

General Columns Advanced Constraints Partitions Parameters Security SQL

Name: estado

Owner: postgres

Schema: rh

Tablespace: Select an item...

Partitioned table?: No

Comment:

Create - Table

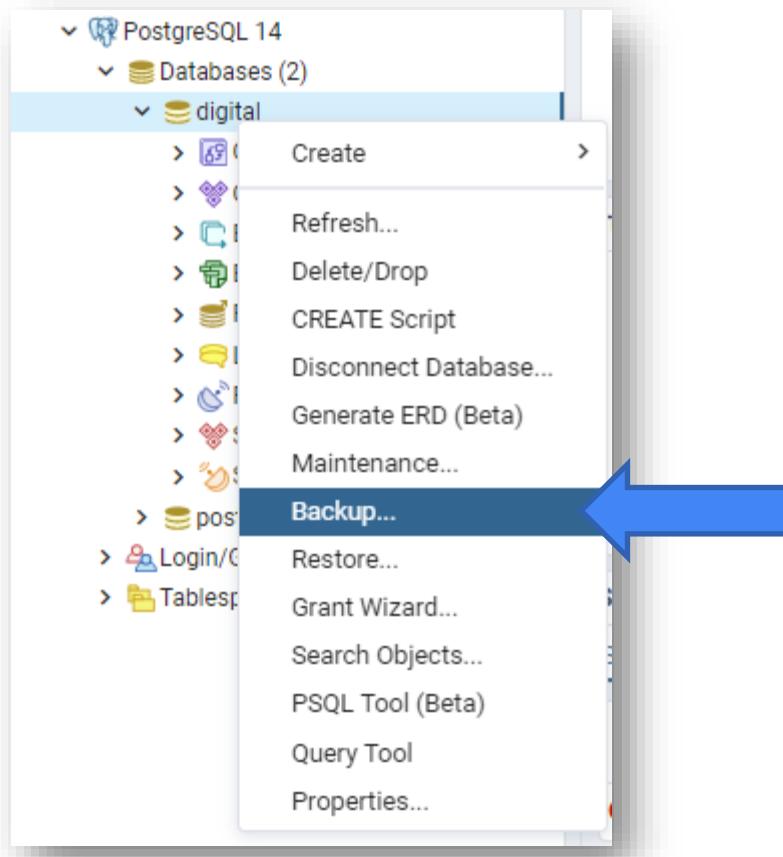
General Columns Advanced Constraints Partitions Parameters Security SQL

Inherited from table(s): Select to inherit from...

Columns

	Name	Data type	Length/Precision	Scale	Not NULL?	Primary key?
<input type="button" value="Edit"/> <input type="button" value="Delete"/>	id	serial			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="button" value="Edit"/> <input type="button" value="Delete"/>	sigla	character varying			<input type="checkbox"/>	<input type="checkbox"/>
<input type="button" value="Edit"/> <input type="button" value="Delete"/>	nome	character varying			<input type="checkbox"/>	<input type="checkbox"/>

Backup e Restore



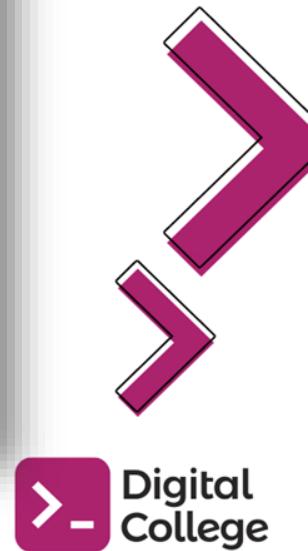
Backup e Restore

Backup (Database: digital)

General Dump options

Filename	c:\temp\digital_backup	...
Format	Custom	▼
Compression ratio		
Encoding	Select an item...	▼
Number of jobs		
Role name	Select an item...	▼

i **?** **Cancel** **Backup**



Backup e Restore

Backup (Database: digital)

General [Dump options](#)

Sections

Pre-data	<input type="checkbox"/> No	Data	<input type="checkbox"/> No
Post-data	<input type="checkbox"/> No		

Type of objects

Only data	<input type="checkbox"/> No	Only schema	<input type="checkbox"/> No
Blobs	<input checked="" type="checkbox"/> Yes		

Do not save

Owner	<input type="checkbox"/> No	Privilege	<input type="checkbox"/> No
Tablespace	<input type="checkbox"/> No	Unlogged table data	<input type="checkbox"/> No
Comments	<input type="checkbox"/> No		

Queries

Use Column Inserts	<input type="checkbox"/> No	Use Insert Commands	<input type="checkbox"/> No
Include CREATE DATABASE statement	<input type="checkbox"/> No	Include DROP DATABASE statement	<input type="checkbox"/> No
Load Via Partition Root	<input type="checkbox"/> No		

Disable

Trigger	<input type="checkbox"/> No	\$ quoting	<input type="checkbox"/> No
---------	-----------------------------	------------	-----------------------------

Miscellaneous

With OID(s)	<input type="checkbox"/> No	Verbose messages	<input checked="" type="checkbox"/> Yes
Force double quote on identifiers	<input type="checkbox"/> No	Use SET SESSION AUTHORIZATION	<input type="checkbox"/> No

[i](#) [?](#) [Cancel](#) [Backup](#)

Backup e Restore

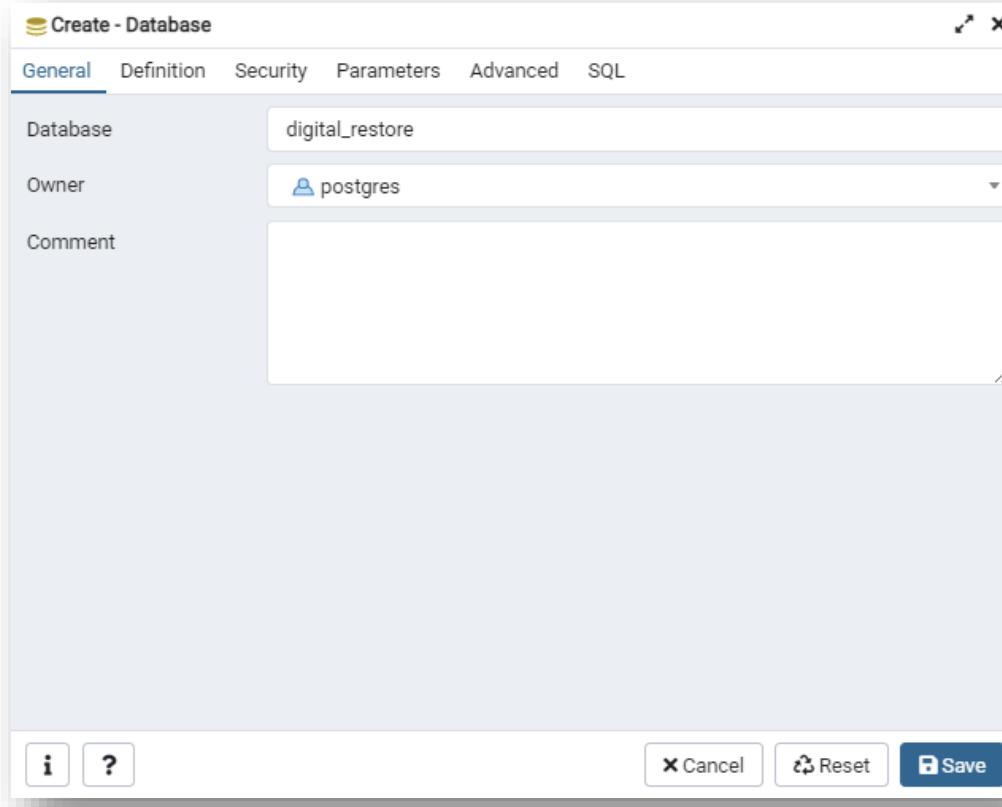
Backup (Database: digital)

General Dump options

Filename	c:\temp\digital_backup	...
Format	Custom	▼
Compression ratio		
Encoding	Select an item...	▼
Number of jobs		
Role name	Select an item...	▼

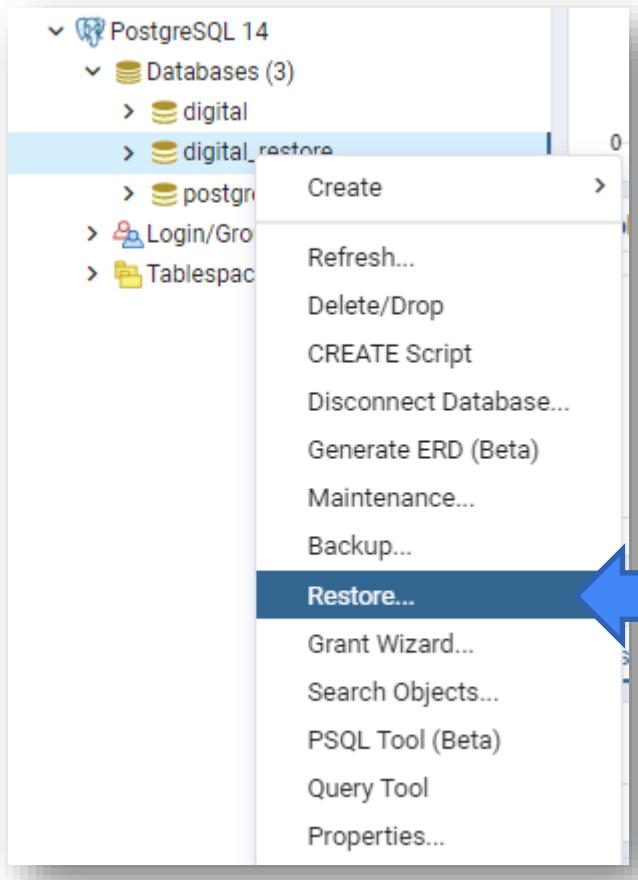
i **?** **Cancel** **Backup**

Backup e Restore



Criar um banco de dados
chamado digital_restore

Backup e Restore



Botão direito em cima do banco de dados

Backup e Restore

Restore (Database: digital_restore)

General Restore options

Format: Custom or tar

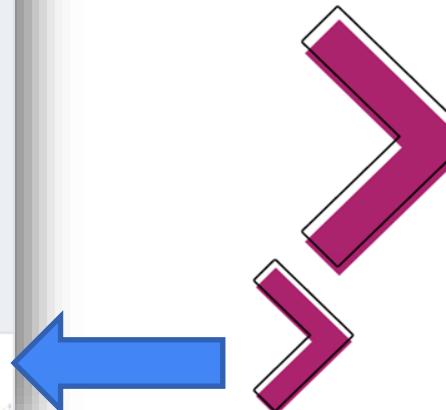
Filename: C:\temp\digital_backup ...

Number of jobs:

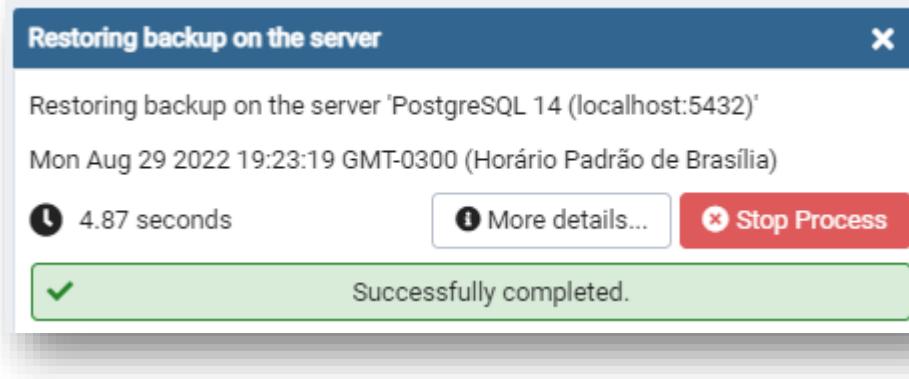
Role name: Select an item...

i ?

X Cancel Restore



Backup e Restore



Backup e Restore

- ›  digital_restore
 - ›  Casts
 - ›  Catalogs
 - ›  Event Triggers
 - ›  Extensions
 - ›  Foreign Data Wrappers
 - ›  Languages
 - ›  Publications
 - ›  Schemas (4)
 - ›  public
 - ›  rh
 - ›  vendas
 - ›  vendas2
 - ›  Subscriptions

Banco de Dados restaurado

Desafio

Consultando e calculando indicadores



Elabore consultas para os indicadores e solicitações abaixo:

1. Apresente a *query* para listar o código e o nome do vendedor com maior número de vendas (*contagem*), e que estas vendas estejam com o status concluída. As colunas presentes no resultado devem ser, portanto, **codigovendedor** (cdvdd) e **nomevendedor** (nmvdd).
2. Apresente a *query* para listar o código e nome do produto mais vendido entre as datas de 2014-02-03 até 2018-02-02. As colunas presentes no resultado devem ser **codigoproduto** (cdpro) e **nomeproduto** (nmpro).
3. Apresente a *query* para listar o código e nome cliente com maior gasto na loja. As colunas presentes no resultado devem ser **codigocliente** (cdcli), **nomecliente** (nmcli) e **gasto**, esta última representando o somatório das vendas atribuídas ao cliente.
4. Apresente a *query* para listar código, nome e data de nascimento dos dependentes do vendedor com menor valor total bruto em vendas (não sendo zero). As colunas presentes no resultado devem ser **codigodependente** (cddep), **nomedependente** (nmdep), **datanascimento** (dtnasc).

**todas as consultas devem trazer apenas vendas com status concluído.*

Elabore consultas para os indicadores e solicitações abaixo:

5. Apresente a *query* para listar os 3 produtos menos vendidos pelos canais de E-Commerce ou Matriz. As colunas presentes no resultado devem ser canalvendas (canal), codigoproduto (cdpro), nomeproduto (nmpro) e quantidade_vendas.
6. Apresente a *query* para listar o gasto médio por estado da federação. As colunas presentes no resultado devem ser estado e gastomedio. Considere apresentar a coluna gastomedio arredondada na segunda casa decimal.
7. Apresente a *query* para listar o código das vendas (cdven) identificadas como deletadas. Apresente o resultado em ordem crescente.
8. Apresente a *query* para listar a quantidade média vendida de cada produto agrupado por estado da federação. As colunas presentes no resultado devem ser estado e nomeproduto (nmprod) e quantidade_media. Considere arredondar o valor da coluna quantidade_media na quarta casa decimal. Ordene os resultados pelo estado (1º) e nome do produto (2º).
9. Calcule a receita bruta por ano.
10. Calcule a receita bruta por ano e por estado.
11. Proponha um indicador...

*todas as consultas devem trazer apenas vendas com status concluído.

Recados

1. *Apresentação de Trabalhos da Unidade 1 (Aldeota) – Segunda as 18hs (estão convidados, lembrando que temos aula aqui!)*
2. *Imersão – Dia a Dia do Analista de Dados (27/04 as 19hs – Sul)*
3. *Atividades – Todas as Atividades, Desafios que Eu solicitar a partir de hoje, deverá ser entregue pelo Discord (inbox para mim)*



The banner features the Digital College logo and address (Av. Washington Soares, 3663 – 4º andar, Torre 2 – Fortaleza/CE). It highlights the date (27 ABR 2023), time (19H ÀS 22H), and theme (DIA A DIA DE UM ANALISTA DE DADOS). Below the banner is a photograph of two people working at a computer displaying a 'Financial Executive Dashboard' with various charts and graphs.

Obrigad@!



Digital
College

ENSINO DE HABILIDADES DIGITAIS

digitalcollege.com.br • [@digitalcollegebr](https://twitter.com/digitalcollegebr)

Módulo 2: Bancos de Dados Dimensionais

