



Data Science  
Academy

Data Science Academy [caiodantas@bb.com.br](mailto:caiodantas@bb.com.br) 5de05db6e32fc31c847c6792



# Big Data Analytics com R e Microsoft Azure Machine Learning



Data Science  
Academy

Data Science Academy [caiodantas@bb.com.br](mailto:caiodantas@bb.com.br) 5de05db6e32fc31c847c6792

# Big Data Analytics com R e Microsoft Azure Machine Learning

## Trabalhando com Bancos de Dados Relacionais e NoSQL em R

### Seja Bem-Vindo(a)!



Data Science  
Academy

Data Science Academy [caiodantas@bb.com.br](mailto:caiodantas@bb.com.br) 5de05db6e32fc31c847c6792

# Trabalhando com Bancos de Dados Relacionais e NoSQL em R





# Trabalhando com Bancos de Dados Relacionais e NoSQL em R



## **Bancos de Dados Como Fonte de Dados**

Bancos de Dados Relacionais – Dados Estruturados  
(RDBMS – Relational Database Management System)

Bancos de Dados Não Relacionais – Dados Semi ou  
Não Estruturados (NoSQL)



# Trabalhando com Bancos de Dados Relacionais e NoSQL em R



**Teremos muitas configurações de Sistema Operacional.**

**Saber configurar e ajustar o SO é tarefa importante para montar seu ambiente de desenvolvimento em análise de dados.**



Data Science  
Academy

Data Science Academy [caiodantas@bb.com.br](mailto:caiodantas@bb.com.br) 5de05db6e32fc31c847c6792

# Big Data Analytics com R e Microsoft Azure Machine Learning

## Introdução aos Bancos de Dados Relacionais

Seja Bem-Vindo(a)!



# Introdução aos Bancos de Dados Relacionais

## Exemplo de um BD Relacional

Empregado

NumEmp	NomeEmp	Salário	Dept
032	J Silva	380	21
074	M Reis	400	25
089	C Melo	520	28
092	R Silva	480	25
112	R Pinto	390	21
121	V Simão	905	28
130	J Neves	640	28

Departamento

NumDept	NomeDept	Ramal
21	Pessoal	142
25	Financeiro	143
28	Técnico	144





# Introdução aos Bancos de Dados Relacionais



Um banco de dados relacional é um conjunto de arquivos físicos que armazenam os dados em formato estruturado, organizados em tabelas, que por sua vez são divididas em linhas e colunas.

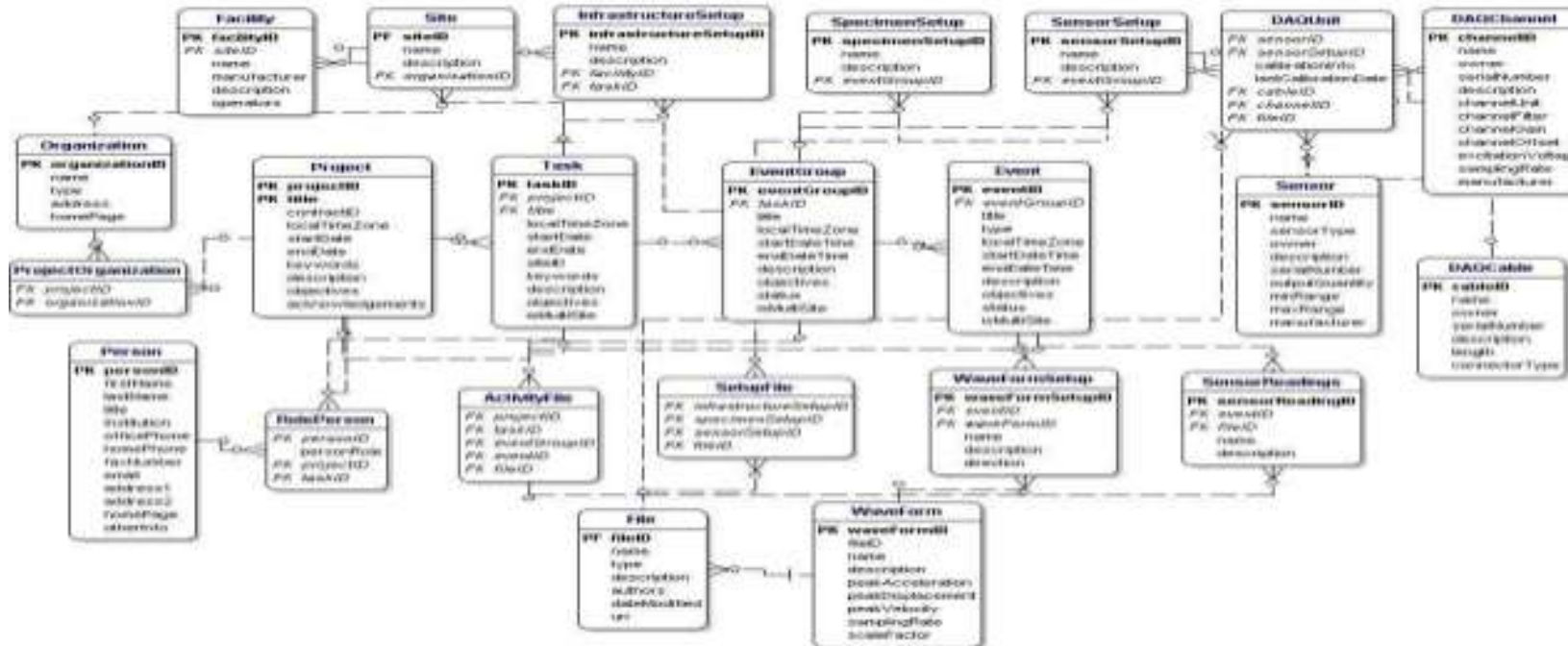
Usamos um SGBD (Sistema Gerenciador de Bancos de Dados ou em inglês RDBMS – Relational Database Management System) para gerenciar os arquivos do banco de dados.





# Introdução aos Bancos de Dados Relacionais

## MODELAGEM RELACIONAL

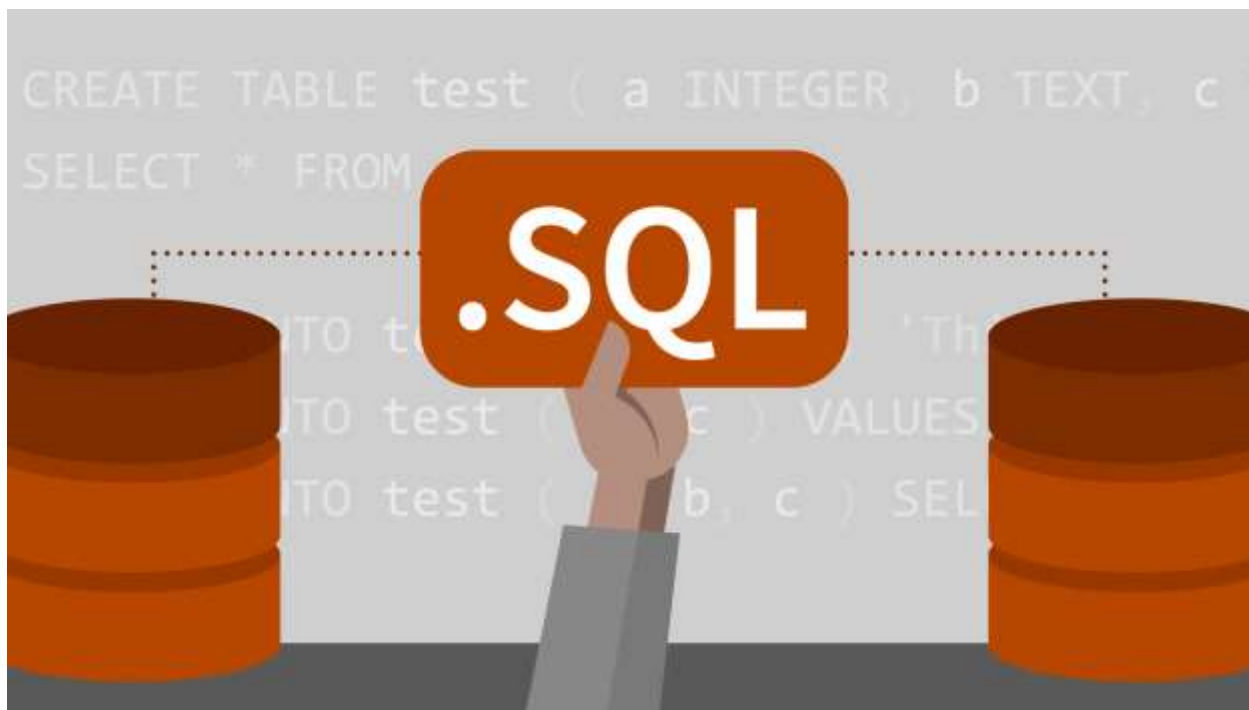


A criação de um banco de dados relacional é baseada na modelagem relacional, trabalho normalmente feito por analistas de sistemas ou administradores de bancos de dados.



# Introdução aos Bancos de Dados Relacionais

E usamos linguagem SQL (Structured Query Language) para manipular os dados em um banco de dados relacional.





# Introdução aos Bancos de Dados Relacionais



A Linguagem R suporta diversos bancos de dados relacionais através de pacotes específicos para cada banco de dados e podemos ainda fazer conexão aos bancos de dados usando um driver padrão.



Data Science  
Academy

Data Science Academy [caiodantas@bb.com.br](mailto:caiodantas@bb.com.br) 5de05db6e32fc31c847c6792

# Introdução a Linguagem SQL

## Linguagem SQL (Structured Query Language)



# Introdução a Linguagem SQL

## Tabela

Coluna ou  
Variável



ID	NOME	IDADE	CIDADE
0001	Pele	120	Roma
0002	Zico	110	Paris
0003	Garrincha	105	Vienna

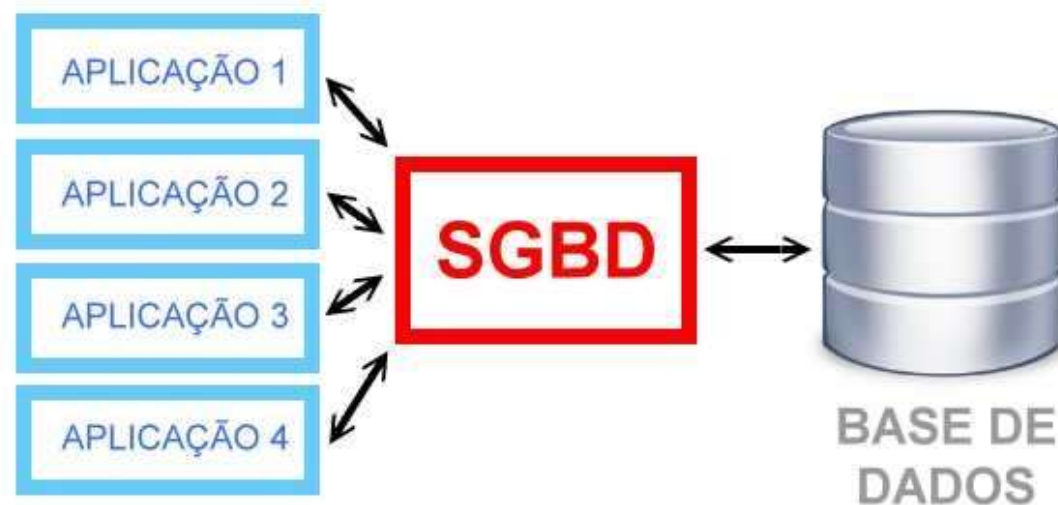
Linha  
ou  
Registro





# Introdução a Linguagem SQL

## Sistemas Gerenciadores de Bancos de Dados Relacionais





# Introdução a Linguagem SQL

## Sistemas Gerenciadores de Bancos de Dados Relacionais

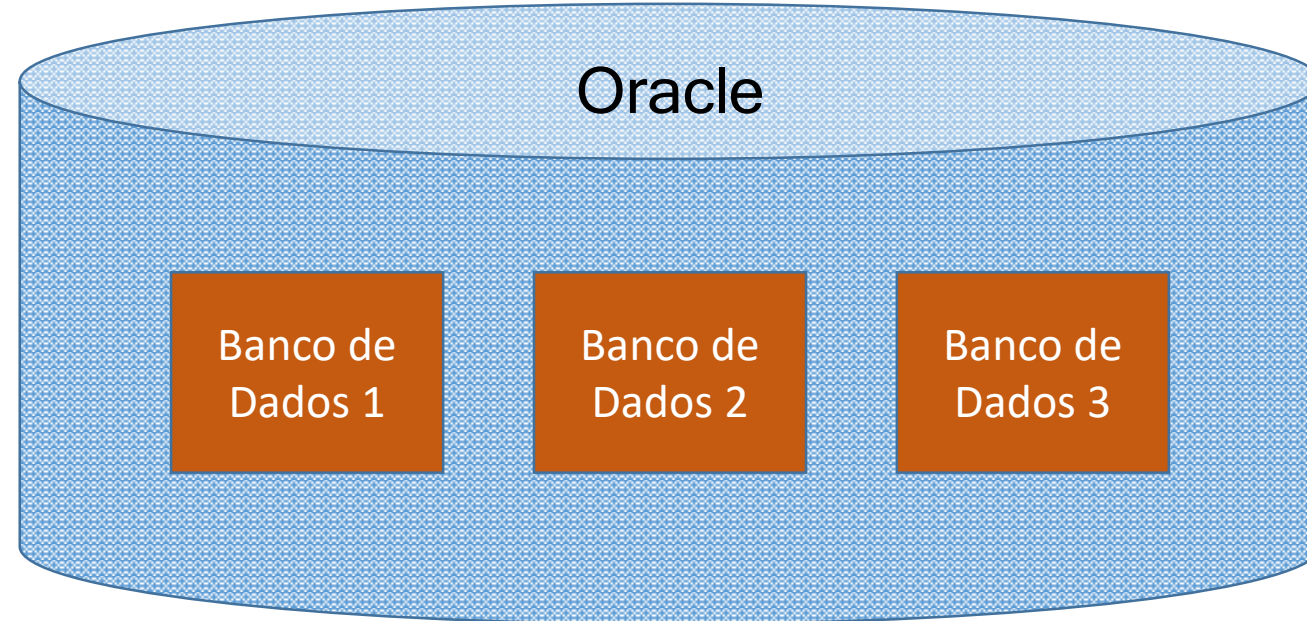
Tudo que fazemos no banco de dados, passa pelo SGBD





# Introdução a Linguagem SQL

## Sistemas Gerenciadores de Bancos de Dados Relacionais

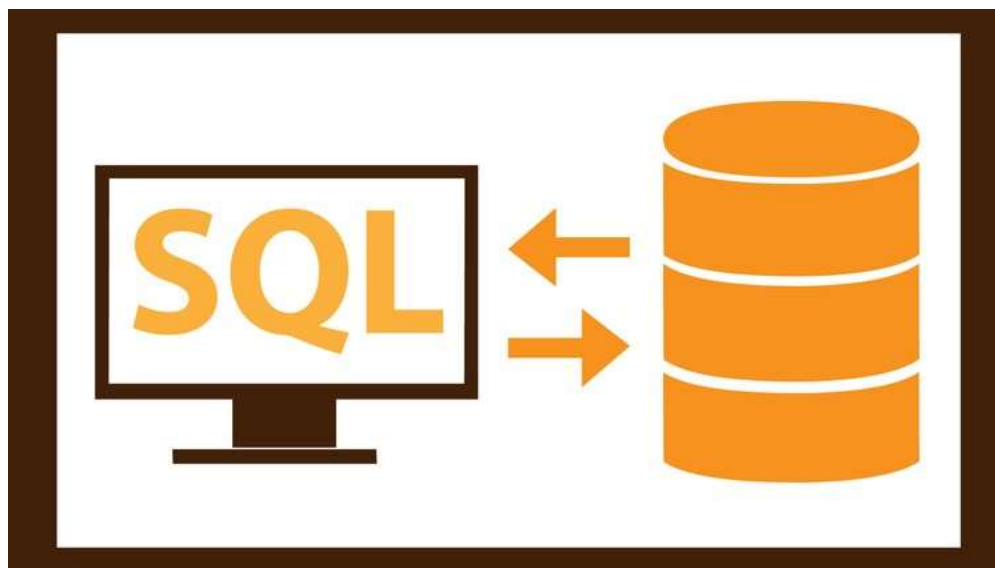




Data Science  
Academy

Data Science Academy caiodantas@bb.com.br 5de05db6e32fc31c847c6792

# Introdução a Linguagem SQL





# Introdução a Linguagem SQL

A Linguagem SQL é implementada de forma diferente em diferentes SGBDs, mas a base da linguagem é a mesma!

Oracle utiliza PL/SQL  
MS Access utiliza o JET SQL  
MS SQL Server utiliza a T-SQL



# Introdução a Linguagem SQL

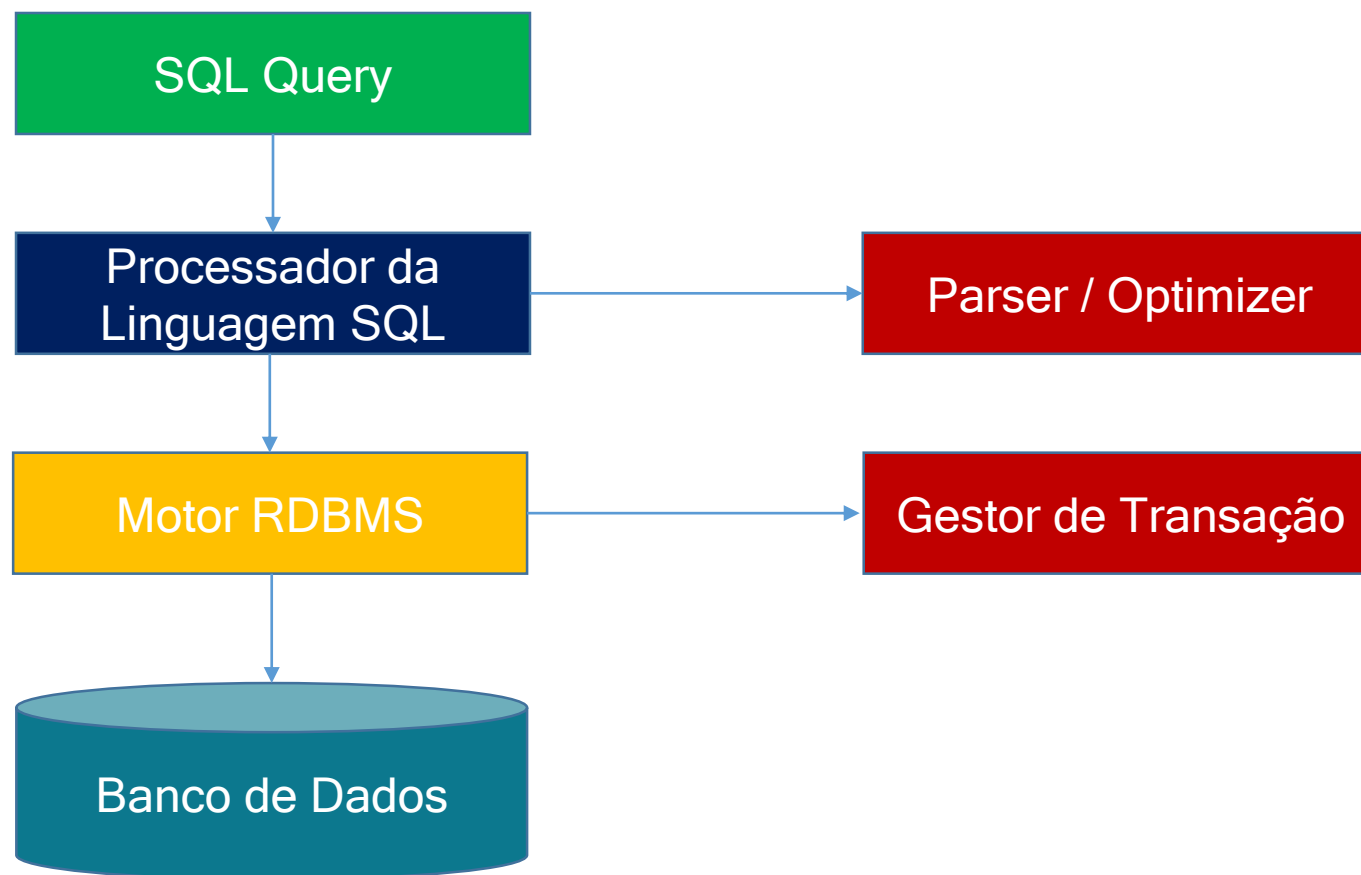
## Quais o benefícios da Linguagem SQL?

- Permite que os usuários acessem dados em sistemas de gerenciamento de bancos de dados relacionais.
- Permite a manipulação de dados armazenados em bancos de dados.
- Permite a criação e remoção de objetos no banco de dados (tabelas, índices, visões, procedimentos armazenados).
- Permite que os usuários possam definir restrições de acesso.



# Introdução a Linguagem SQL

A linguagem SQL é processada pelo engine (motor) do SGBD.





# Introdução a Linguagem SQL

## Existem 3 Tipos de Instrução SQL



DDL



DML



DCL



# Introdução a Linguagem SQL



DDL

## DDL – Data Definition Language

- Create
- Alter
- Drop





# Introdução a Linguagem SQL



DML

## DML – Data Manipulation Language

- Select
- Insert
- Delete
- Update



# Introdução a Linguagem SQL



DCL

## DCL – Data Control Language

- Revoke
- Alter



Data Science  
Academy

Data Science Academy [caiodantas@bb.com.br](mailto:caiodantas@bb.com.br) 5de05db6e32fc31c847c6792

# Introdução aos Bancos de Dados NoSQL

## Bancos de Dados NoSQL (Not Only SQL)





# Introdução aos Bancos de Dados NoSQL

NoSQL é uma tecnologia de banco de dados projetada para suportar os requisitos de aplicações em nuvem e arquitetado para superar em escala e desempenho as limitações de bancos de dados relacionais (RDBMS).



# Introdução aos Bancos de Dados NoSQL

Bancos de dados NoSQL são usados para armazenar dados semi ou não estruturados e ganharam notoriedade por conta do crescimento do Big Data e das aplicações modernas.



Data Science  
Academy

Data Science Academy [caiodantas@bb.com.br](mailto:caiodantas@bb.com.br) 5de05db6e32fc31c847c6792

# Introdução aos Bancos de Dados NoSQL







# Introdução aos Bancos de Dados NoSQL

Os principais bancos de dados NoSQL são:

Graph	Neo4j
	FlockDB
	ArangoDB
	GraphDB

Document	MongoDB
	CouchDB
	RavenDB
	Terrastore

Key-Value	Oracle NoSQL DB
	Memcache
	Redis
	Voldemort

Column	HBase
	Cassandra
	Hypertable
	Accumulo



# Introdução aos Bancos de Dados NoSQL

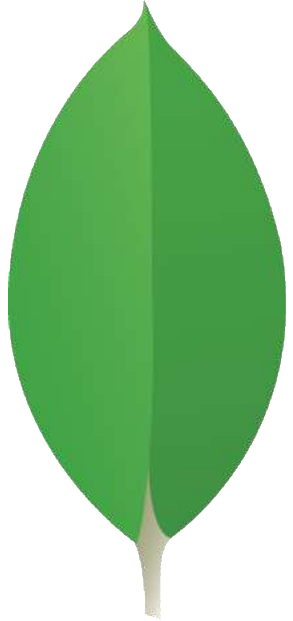
E por que devo aprender a usar um banco de dados  
NoSQL?



Data Science  
Academy

Data Science Academy caiodantas@bb.com.br 5de05db6e32fc31c847c6792

# Introdução aos Bancos de Dados NoSQL



mongoDB®



# Introdução aos Bancos de Dados NoSQL

MongoDB	RDBMS
Database	Database
Collection	Tabela
Document	Linha/Tupla
Field	Coluna
Embedded Documents	Join de Tabelas
Primary Key	Primary Key



# Muito Obrigado por Participar!



Tenha uma Excelente Jornada de Aprendizagem.

Equipe Data Science Academy