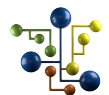
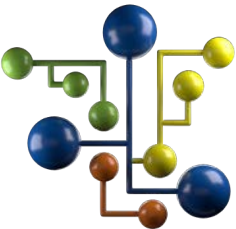


Data Science
Academy

Data Science Academy eduardo.ads1814@gmail.com +354792627416532060075

Big Data Real-Time Analytics com Python e Spark



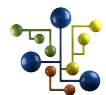
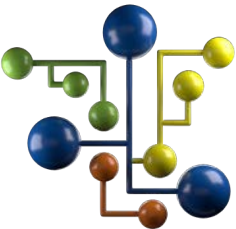


Data Science
Academy

Data Science Academy eduardo.ads1814@gmail.com 63a47926274fe5d3f2060075

Big Data Real-Time Analytics com Python e Spark

Seja muito bem-vindo(a)!

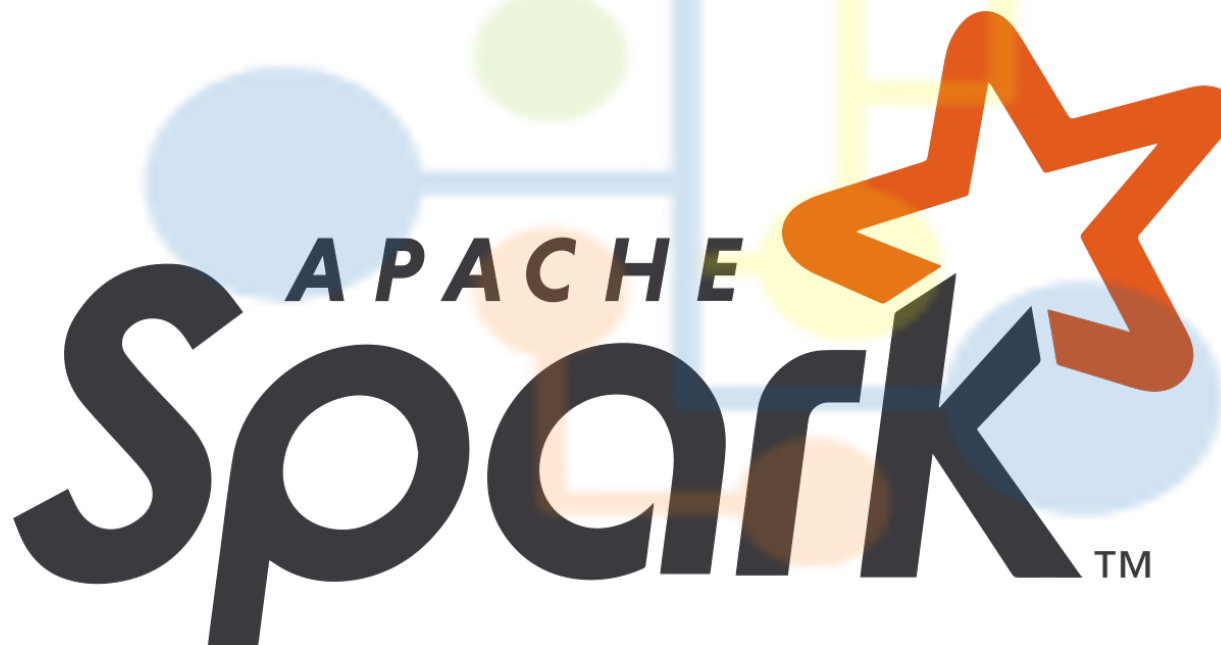


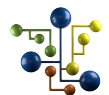
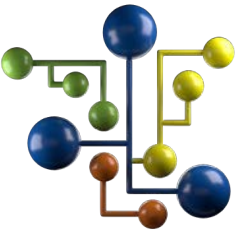
Data Science
Academy

Data Science Academy eduardo.ads1814@gmail.com 63a47926274fe5d3f2060075

Big Data Real-Time Analytics com Python e Spark

Apache Spark SQL





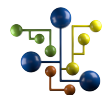
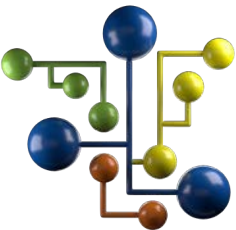
Data Science
Academy

Data Science Academy eduardo.ads1814@gmail.com 63a47926274fe5d3f2060075

Big Data Real-Time Analytics com Python e Spark

Introdução à Linguagem SQL



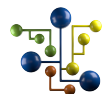
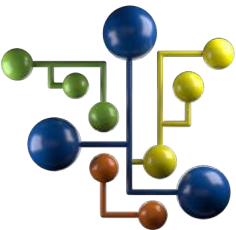


Introdução à Linguagem SQL

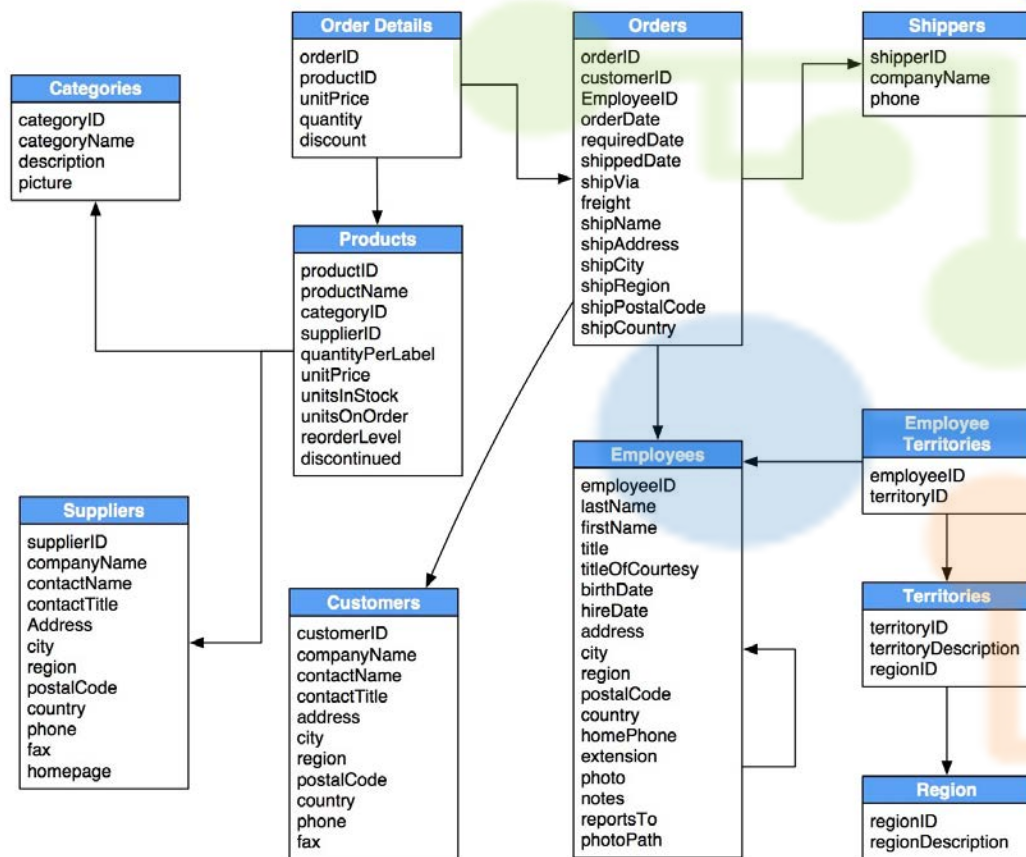
Com SparkSQL nosso principal trabalho é definir as queries SQL.

Logo, conhecer Linguagem SQL é fundamental.

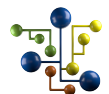
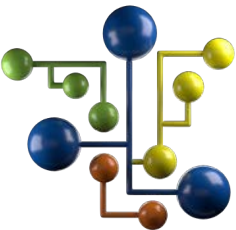
```
results = spark.sql(  
    "SELECT * FROM people")  
names = results.map(lambda p: p.name)
```



Introdução à Linguagem SQL



Usamos Linguagem SQL para consultar e manipular dados em bancos de dados.

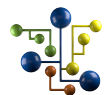
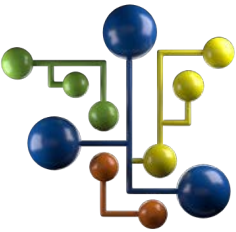


Introdução à Linguagem SQL



SQL

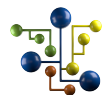
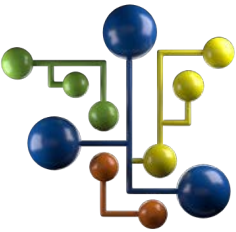
- SQL (ou Structured Query Language) surgiu em meados da década de 70, sendo resultado de um estudo de membros do laboratório de pesquisa da IBM em San Jose, Califórnia.
- Este estudo tinha foco em desenvolver uma linguagem que pudesse ser adaptada ao modelo relacional. O primeiro sistema de banco de dados baseado em SQL tornou-se comercial no final dos anos 70 juntamente com outros sistemas de bancos relacionais.
- O sucesso da linguagem foi tão grande que obrigou o ANSI (American National Standards Institute) a padronizar as implementações da linguagem. O padrão ANSI SQL é o que abordamos aqui no curso.
- Atualmente, a maior parte dos SGBDs (Sistemas Gerenciadores de Bancos de Dados) segue criteriosamente esta padronização, podendo ter algumas variações, mas mesmo assim não afetando a padronização global da linguagem e tornando a portabilidade mais fácil.



Introdução à Linguagem SQL



- Para executar código SQL precisamos de um motor SQL. Os principais SGBDs como Oracle, SQL Server, MySQL, DB2 e PostgreSQL têm seu próprio motor SQL.
- Ferramentas de visualização de dados como Power BI e Tableau permitem o uso de SQL para extrair dados de diferentes fontes.
- Diversos pacotes de manipulação de dados em Python e R são na verdade implementações da Linguagem SQL.
- Podemos ainda executar diretamente a linguagem SQL a partir de diversas linguagens de programação, bem como no SparkSQL.
- A sintaxe da Linguagem SQL pode ser ligeiramente diferente de um SGBD para outro, mas o padrão ANSI é amplamente suportado.

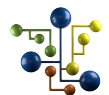
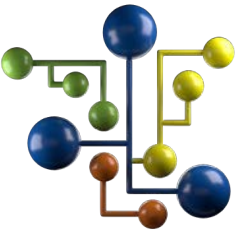


Introdução à Linguagem SQL

As instruções SQL se dividem em 3 tipos principais:

- **Linguagem de Manipulação de Dados ou DML (Data Manipulation Language)**
- **Linguagem de Definição de Dados ou DDL (Data Definition Language)**
- **Linguagem de Controle de Dados ou DCL (Data Control Language)**





Introdução à Linguagem SQL

Instruções DML:

INSERT: Instrução utilizada para inserir registros em uma tabela.

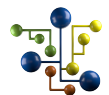
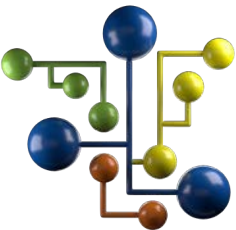
Exemplo: `INSERT INTO tb_cliente(id, nome, sobrenome) VALUES(100, 'Bob', 'Tavares');`

UPDATE: Instrução utilizada para alterar valores de uma ou mais linhas em uma tabela.

Exemplo: `UPDATE tb_cliente SET nome = 'Zico' WHERE id = 100;`

DELETE: Instrução utilizada para excluir um ou mais registros de uma tabela.

Exemplo: `DELETE FROM tb_cliente WHERE id = 100;`



Introdução à Linguagem SQL

Instruções DDL:

CREATE: Instrução utilizada para criar objetos no banco de dados.

Exemplo (criar uma tabela):

```
CREATE TABLE tb_fornecedor (id INT PRIMARY KEY, nome VARCHAR(50));
```

ALTER: Instrução utilizada para alterar a estrutura de um objeto.

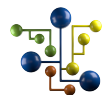
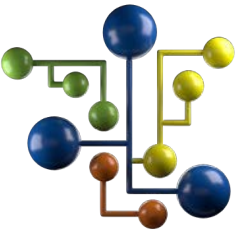
Exemplo (adicionar uma coluna em uma tabela existente):

```
ALTER TABLE tb_fornecedor ADD telefone CHAR(10);
```

DROP: Instrução utilizada para remover um objeto do banco de dados.

Exemplo (remover uma tabela):

```
DROP TABLE tb_fornecedor;
```



Introdução à Linguagem SQL

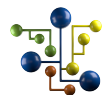
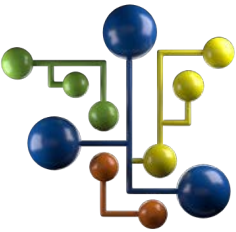
Instruções DCL:

GRANT: Instrução que autoriza um usuário do SGBD a executar alguma operação.
Exemplo (dar permissão de consulta na tabela tb_fornecedor para o usuário maria):

```
GRANT SELECT ON tb_fornecedor TO maria;
```

REVOKE: Instrução que remove a permissão de um usuário executar alguma operação.
Exemplo (não permitir que o usuário maria crie tabelas no banco de dados):

```
REVOKE CREATE TABLE FROM maria;
```

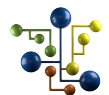
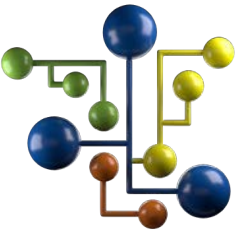



Introdução à Linguagem SQL

Temos ainda mais 2 tipos específicos de instruções:

- **Linguagem de Consultas de Dados ou DQL (Data Query Language)**
- **Linguagem de Controle de Transações (Data Transaction Language)**

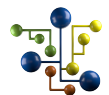
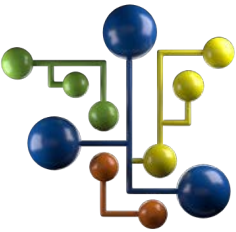




Introdução à Linguagem SQL

Instruções DQL (SELECT):

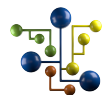
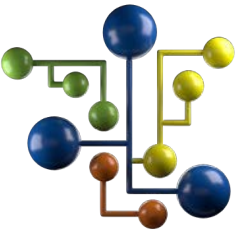
- **Cláusulas** (FROM, WHERE, GROUP BY, HAVING, ORDER BY, DISTINCT)
- **Operadores Lógicos** (AND, OR, NOT)
- **Operadores de Comparação** (<, >, <>, <=, =, >=, BETWEEN, LIKE)
- **Funções de Agregação** (AVG, COUNT, SUM, MIN, MAX)
- **Exemplo:** `SELECT id, nome FROM tb_cliente WHERE id = 100 AND nome = 'Bob';`



Introdução à Linguagem SQL

Instruções DTL:

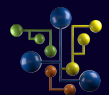
- BEGIN TRANSACTION
- COMMIT
- ROLLBACK



Introdução à Linguagem SQL

Terminologia

- Tabela
- Colunas (atributos)
- Linhas (registros)
- Relação Entre Tabelas
- Chave Primária (Primary Key)
- Chave Estrangeira (Foreign Key)



Data Science
Academy

Data Science Academy eduardo.ads1814@gmail.com 63a47926274fe5d3f2060075

Formação Cientista de Dados 3.0

Big Data Real-Time Analytics com Python e Spark

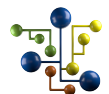
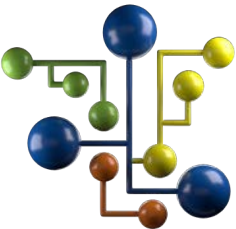
Lab 4

Manipulação de Dados com
SparkSQL, PandaSQL, SQLAlchemy,
PostgreSQL e Docker

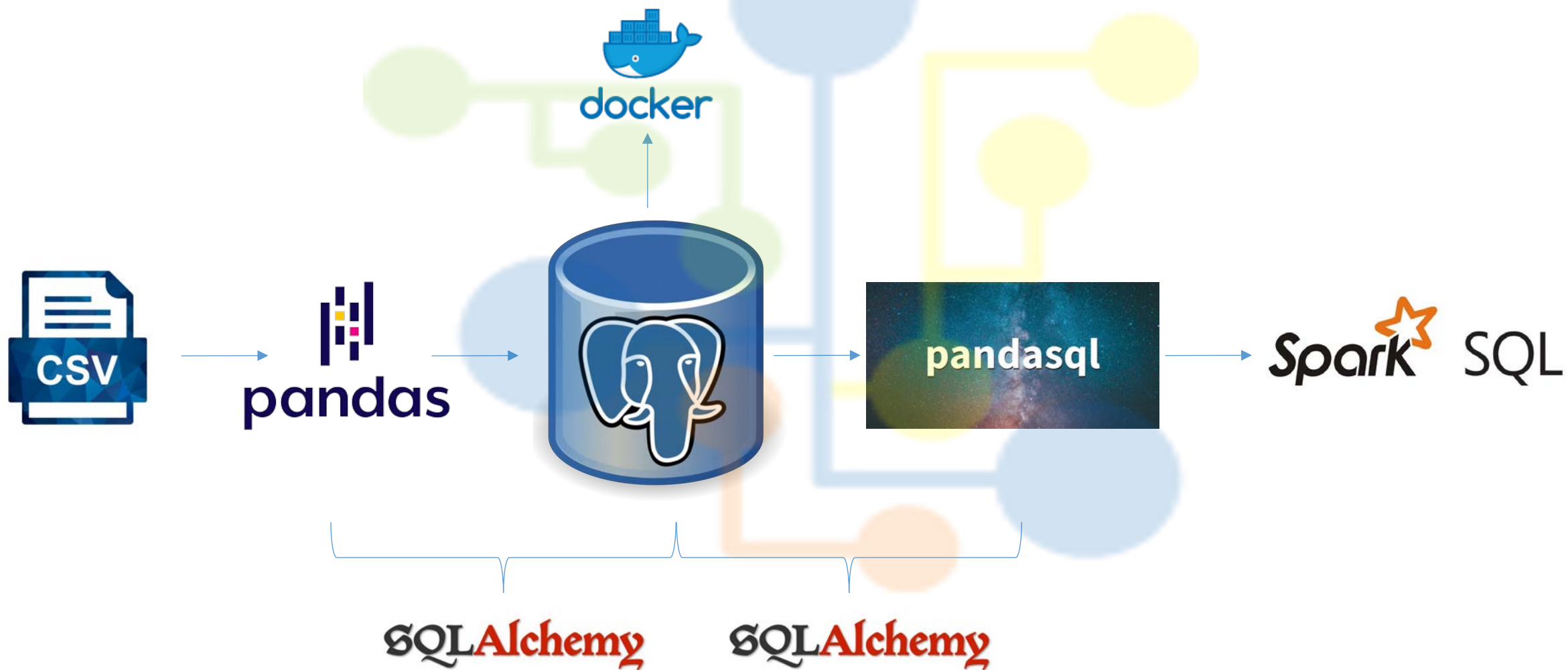
www.datascienceacademy.com.br

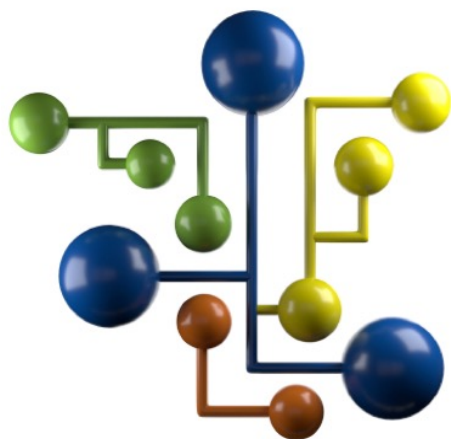
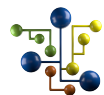


Data Science
Academy



Lab 4 - Visão Geral





Tenha uma Excelente Jornada de Aprendizagem.

Muito Obrigado por Participar!

Equipe Data Science Academy