



**Data Science
Academy**

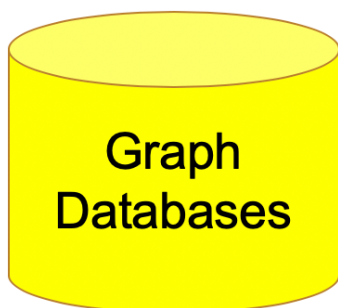
www.datascienceacademy.com.br

Engenharia de Dados com Hadoop e Spark

Categorias de Banco de Dados NoSQL

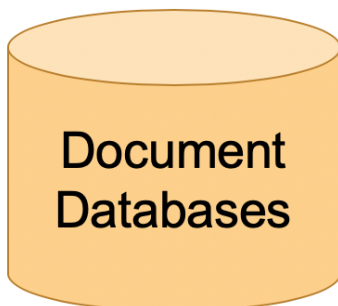
Bancos de Dados NoSQL oferecem 4 categorias principais de bancos de dados:

- Graph databases
- Document databases
- Key-values stores
- Column family stores



Graph	Neo4J
	FlockDB
	GraphDB
	ArangoDB

Esta categoria de Bancos de Dados NoSQL, geralmente é aderente a cenários de rede social on-line, onde os nós representam as entidades e os laços representam as interconexões entre elas. Desta forma, é possível atravessar o gráfico seguindo as relações. Esta categoria tem sido usada para lidar com problemas relacionados a sistemas de recomendação e listas de controle de acesso, fazendo uso de sua capacidade de lidar com dados altamente interligados.



Document	MongoDB
	CouchDB
	RavenDB
	Terrastore

Esta categoria de Bancos de Dados NoSQL permite o armazenamento de milhões de documentos. Por exemplo, você pode armazenar detalhes sobre um empregado, junto com o currículo dele (como um documento) e então pesquisar sobre potenciais candidatos a uma vaga, usando um campo específico, como telefone ou conhecimento em uma tecnologia.



Key-value	Oracle NoSQL DB
	MemcacheDB
	Redis
	Amazon DynamoDB

Nesta categoria, os dados são armazenados no formato key-value (chave-valor) e os valores (dados) são identificados pelas chaves. É possível armazenar bilhões de registros de forma eficiente e o processo de escrita é bem rápido. Os dados podem ser então pesquisados através das chaves associadas.



Column	HBase
	Cassandra*
	Hypertable
	Accumulo

* Cassandra é híbrido, Column e Key-value

Também chamados bancos de dados orientados a coluna, os dados são organizados em grupos de colunas e tanto o armazenamento, quando as pesquisas de dados são baseadas em chaves. HBase é um banco de dados NoSQL orientado a famílias de colunas. Mais a frente veremos o que é esse conceito e como criar famílias de colunas em tabelas do HBase.