

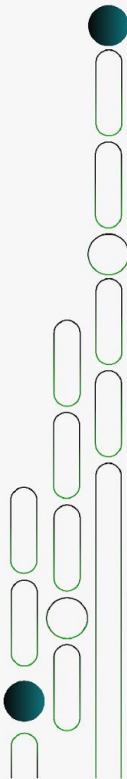


Aprenda com quem faz

> Fundamentos de Bancos de Dados

Segunda Aula Interativa

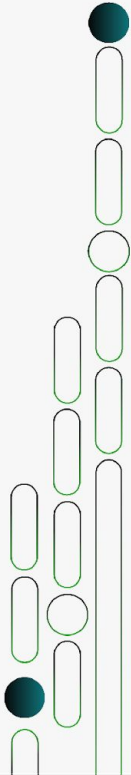
Prof. Diego Bernardes





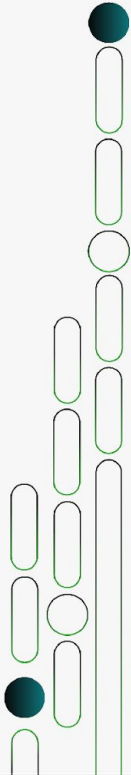
Nesta aula

- Revisão – Auto Relacionamento.
- Bancos de Dados Não Relacionais.
- Discussões e dúvidas.



Auto Relacionamento

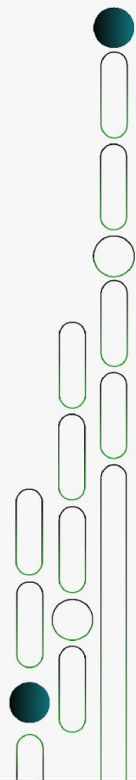
- Considere o requisito do modelo de Dados de Recursos Humanos utilizado no curso:
 - Os funcionários são hierarquicamente organizados de modo a ter um gestor direto, mesmo que este seja de departamentos diferentes.
 - O modelo deve permitir a identificação dos gestores.





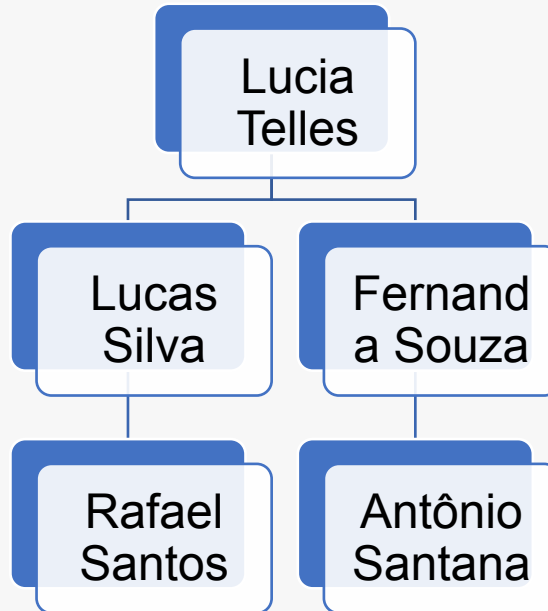
Auto Relacionamento

- Como exemplo, vamos pegar um “recorte” dos seguintes funcionários:
 - Rafael Santos.
 - Lucas Silva.
 - Fernanda Souza.
 - Lucia Telles.
 - Antônio Santana.



Auto Relacionamento

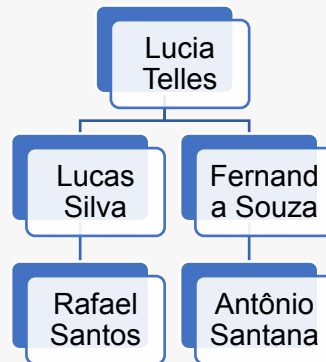
Hierarquia da Empresa



Auto Relacionamento

Tabela: Empregados

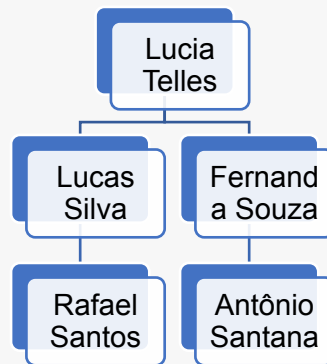
ID	NOME	SALARIO	DEP_ID
1	Lucia Telles	10000	10
2	Fernanda Souza	8000	20
3	Lucas Silva	8000	30
4	Rafael Santos	5000	30
5	Antônio Santana	5000	20



Auto Relacionamento

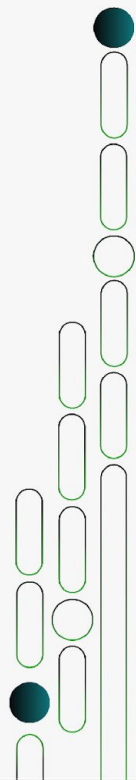
Incluindo gerentes

ID	NOME	SALARIO	DEP_ID	GERENTE_ID
1	Lucia Telles	10000	10	NULL
2	Fernanda Souza	8000	20	1
3	Lucas Silva	8000	30	1
4	Rafael Santos	5000	30	3
5	Antônio Santana	5000	20	2



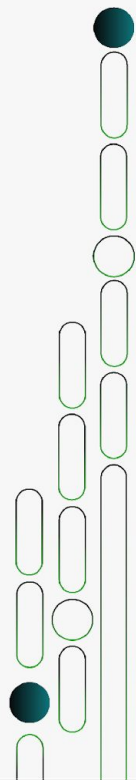


Demonstração Prática





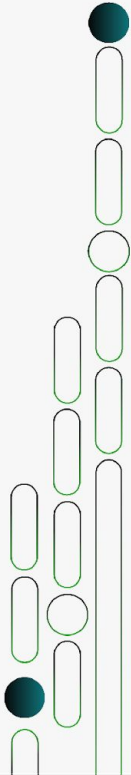
Dúvidas do Chat





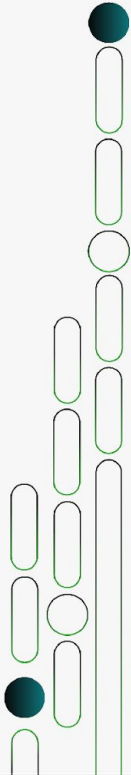
Bancos de Dados NoSQL

- Características principais:
 - Alto volume de dados.
 - Foco em desempenho.
 - Flexibilidade no schema.
 - Restrições controladas pelas aplicações.



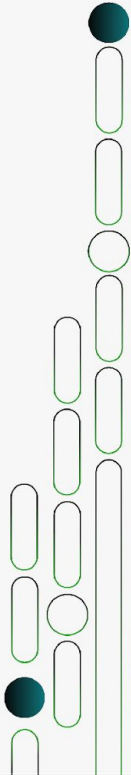
Bancos de Dados NoSQL

- Características principais:
 - 1º V: Volume.
 - Até 2020, se estima que o volume de dados produzidos no mundo chegue a 44 zetabytes (ou o mesmo que 44 trilhões de gigabytes).



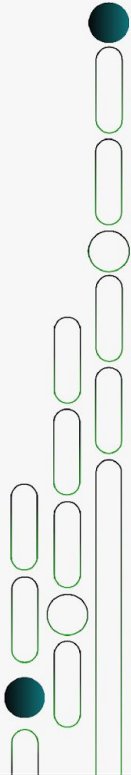
Bancos de Dados NoSQL

- Características principais:
 - 2º V: Velocidade.
 - Para esse grande volume de dados, o objetivo é recuperar o dado em poucos segundos, em “real time” ou “near real time”.



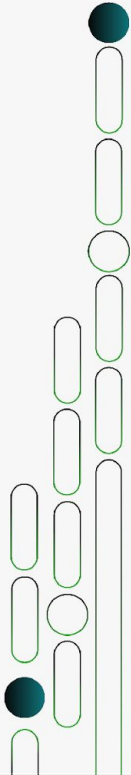
Bancos de Dados NoSQL

- Características principais:
 - 3º V: Variedade.
 - Os dados podem vir de várias fontes de dados e também em diferentes formatos, muitas vezes não estruturados ou semiestruturados.



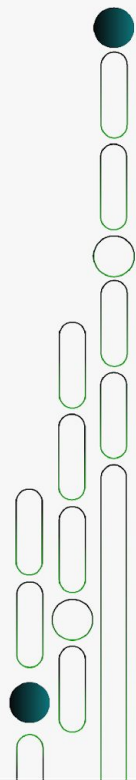
Bancos de Dados NoSQL

- Características principais:
 - 4º V: Veracidade.
 - Diante do grande volume e de diferentes fontes de dados, é importante conseguir distinguir dados que representam a realidade, sem “falsos positivos”.

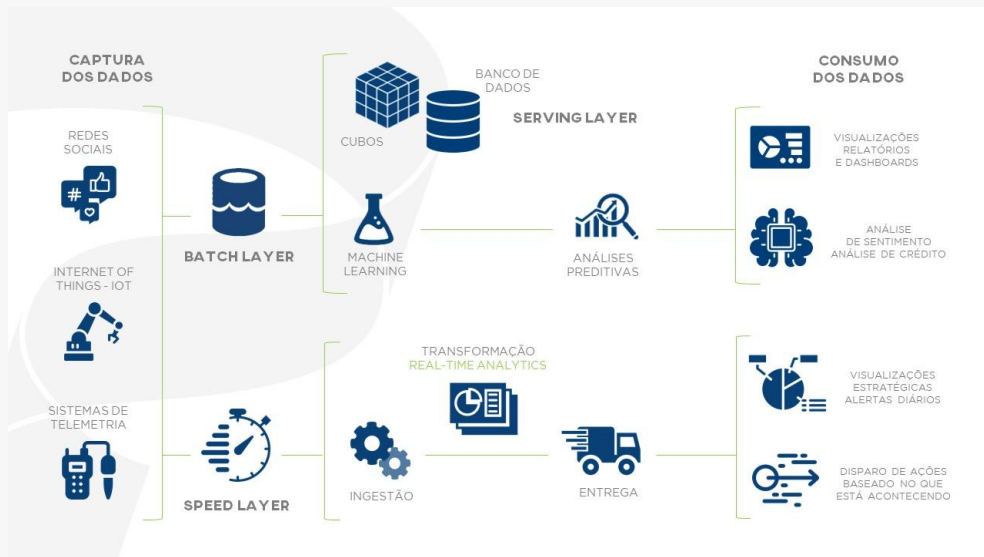


Bancos de Dados NoSQL

- Características principais:
 - 5º V: Valor.
 - Toda essa volumetria deve ser utilizada para trazer valor para o negócio, analisar, diagnosticar, prever e estabelecer tendências.

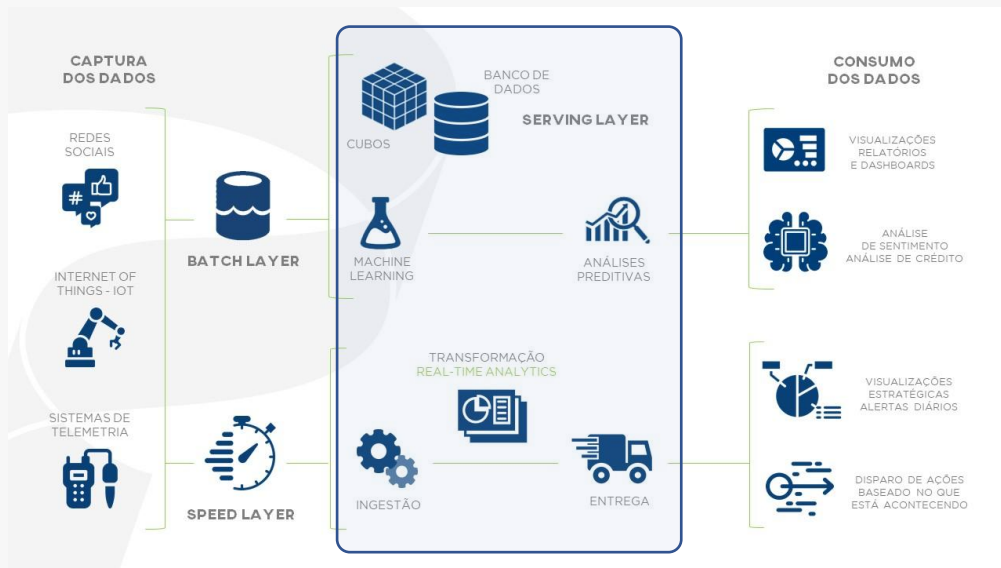


Arquitetura de Dados



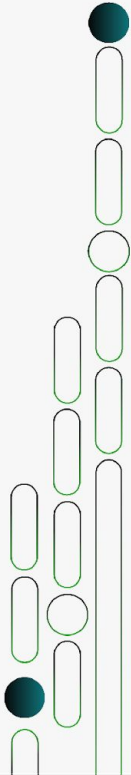


Arquitetura de Dados



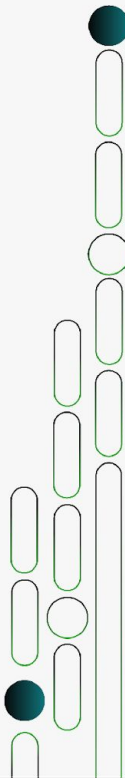
Desafio

- O desafio consistiu na elaboração de dois modelos de dados NoSQL:
 - 01 modelo de dados columnar.
 - 01 modelo de dados de documentos.
- Após a elaboração dos dois modelos, temos subsídios para responder às questões do desafio.



Desafio – Projeto 01

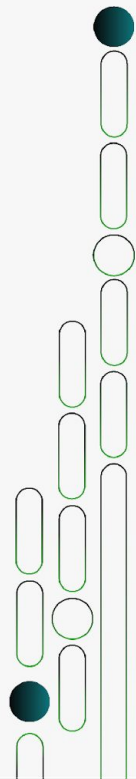
- Criar um sistema de envio e recebimento de notas fiscais da empresa, cuja necessidade de gravação é sem transformar a estrutura original em json para outro formato e que seja possível recuperar as notas através de pesquisas por sua descrição ou produto, ou seja, todo o conteúdo da nota sempre será recuperado quando as consultas são realizadas. Para esse requisito, identificou-se uma probabilidade de modelo de dados orientado a documento, para armazenar os arquivos json e pesquisá-los.





Desafio – Projeto 01

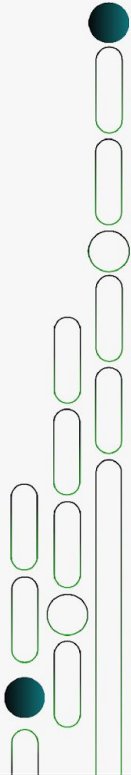
- Você, como analista, deve analisar essa demanda, considerando as seguintes informações de nota que devem ser armazenadas:
 - CNPJ
 - Valor de cada item da nota
 - Fornecedor
 - Quantidade de cada item da nota
 - Data de emissão
 - Descrição de cada item da nota
 - Valor total da nota
 - Nome de cada item da nota





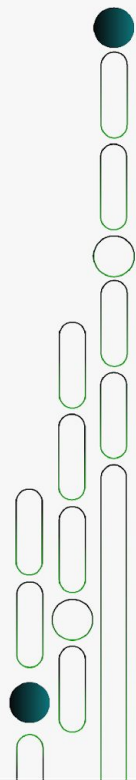
Desafio – Projeto 01

- Lembrando que cada nota contém no mínimo um item, mas pode conter vários. Desenvolva o modelo adequado a essa demanda e responda às perguntas a seguir.



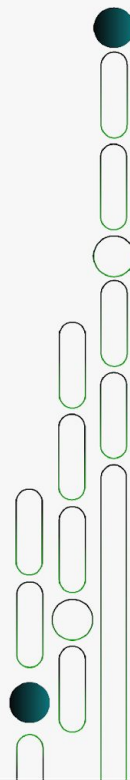


Demonstração Prática



Desafio – Projeto 02

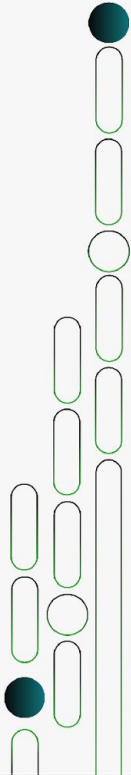
- Criar um sistema analítico, com massivo processamento de consultas aos dados, com poucos momentos de inserção e sem atualizações nos dados, ou seja, um sistema de relatórios. Sabe-se que a carga de dados será responsabilidade de outro setor, mas considerando que os relatórios são distintos entre si, existe uma sinalização de um modelo de dados colunar para atendimento dessa demanda. Sua atuação como analista é descrever esse modelo de modo a permitir o melhor desempenho possível das consultas.





Desafio – Projeto 02

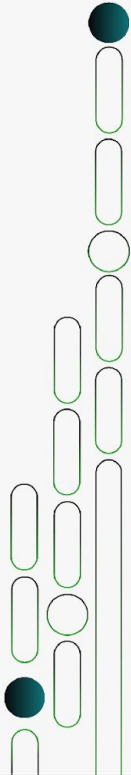
- Regras:
 - Em vendas, os dados utilizados em conjunto são valor e nome do produto, bem como filial e data de venda.
 - Em compras, o valor da compra é exibido sempre junto com o nome do produto, bem como data da venda e comprador.
 - Em propagandas, é importante exibir nome da campanha e público alvo sempre juntos.





Desafio – Projeto 02

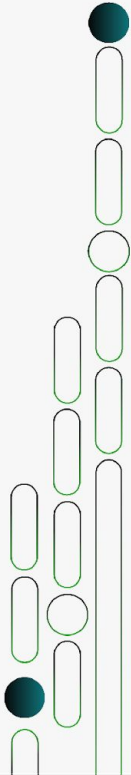
- Informações:
 - Vendas:
 - Valor de uma venda;
 - Nome do produto vendido;
 - Data da venda;
 - Filial.





Desafio – Projeto 02

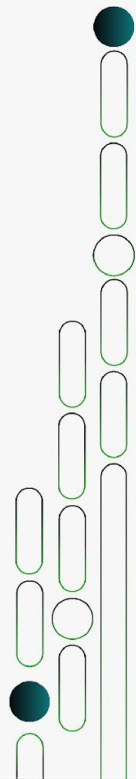
- Informações:
 - Compras:
 - Valor de uma compra;
 - Produto comprado;
 - Nome do vendedor;
 - Nome do comprador.





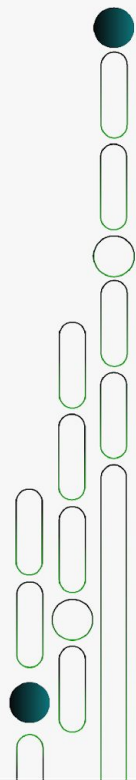
Desafio – Projeto 02

- Informações:
 - Propagandas:
 - Nome da campanha.
 - Data de início da campanha.
 - Data de finalização da campanha.
 - Público-alvo.



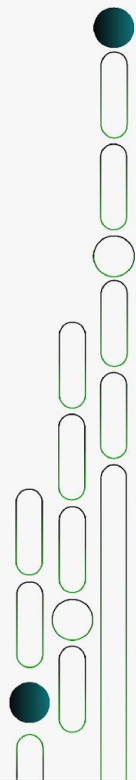


Demonstração Prática





Dúvidas do Chat



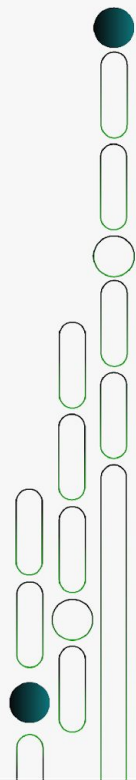


Qual escolher?



Critérios

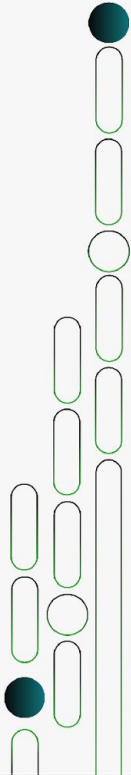
- Origem dos dados:
 - Como o dado entra no sistema?
- Restrições que devem ser aplicadas:
 - Existem regras de nulidade ou dependência?
- Estrutura dos dados:
 - São dados tabulares? Arquivos? Sensores?





Critérios

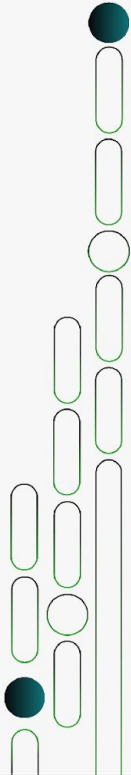
- Estrutura dos dados:
 - Áudio? Vídeo?
- Forma de acesso:
 - Muitas consultas?
 - Muitas gravações?
 - Muitas atualizações?





Critérios

- Suporte constante é um critério importante?
- Segurança x Desempenho.
- Nuvem x On Premises?



Exemplos de Dados

- Documentos

Formatted JSON Data

```
[{"partidas": [{"partida_id": 221469, "clube_casa_id": 283, "clube_casa_posicao": 10, "clube_visitante_id": 289, "aproveitamento_mandante": [{"v", "d"}]}
```

Formatted JSON Data

```
[{"partidas": [{"partida_id": 221469, "clube_casa_id": 283, "clube_casa_posicao": 10, "clube_visitante_id": 289, "aproveitamento_mandante": [{"v", "d"}]}
```

Formatted JSON Data

```
[{"partidas": [{"partida_id": 221469, "clube_casa_id": 283, "clube_casa_posicao": 10, "clube_visitante_id": 289, "aproveitamento_mandante": [{"v", "d"}]}
```

Formatted JSON Data

```
[{"partidas": [{"partida_id": 221469, "clube_casa_id": 283, "clube_casa_posicao": 10, "clube_visitante_id": 289, "aproveitamento_mandante": [{"v", "d"}]}
```

Exemplos de Dados

- Famílias de Colunas

Person column family

joe.doe@gmail.com

firstName	Joe
lastName	Doe
URL	www.joe.doe.com
university	Paris8
type	User, Author, ConfMember, Reviewer

marie.smith@yahoo.com

firstName	Marie
lastName	Smith
URL	www.msmith.org
university	MIT
type	User, ConfMember

Paper column family

Paper202

title	...
abstract	...
content	...
authors	joe.doe@gmail.com
submissionDate	5/5/2011

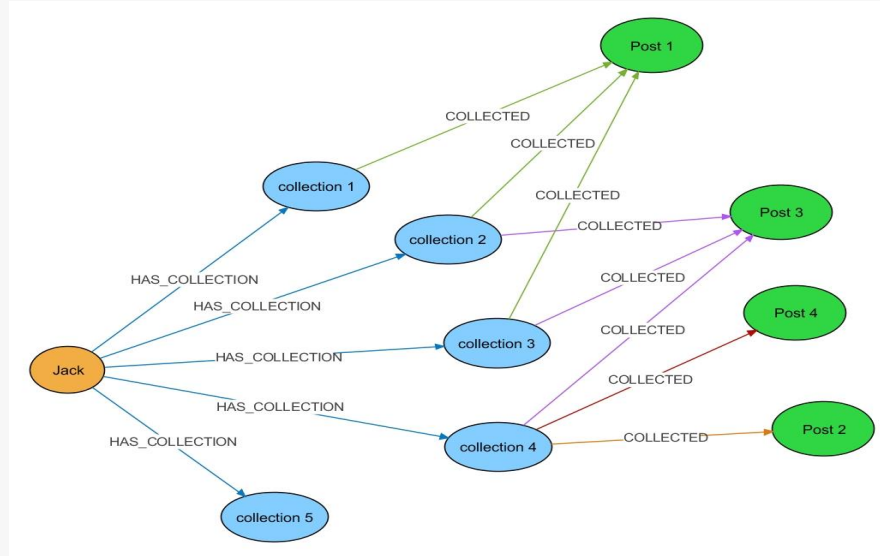
Review column family

Review66

paper	Paper305
abstract	...
content	...
submissionDate	5/15/2011
author	joe.doe@gmail.com

Exemplos de Dados

- Grafos





Conclusão

- Na aula de hoje, fizemos uma breve revisão do conteúdo e pudemos discutir itens adicionais sobre modelagem de dados NoSQL.

