

A vertical bar on the left side of the slide, transitioning from green at the top to blue at the bottom.

Uso dos Dados na tomada de decisão



HENRIQUE SANTANA
Head of Data & Machine Learning
@ **Cloud Humans**



AGENDA

- **Não temos dados**
- **Agora temos dados**
- **O que fazer com os dados**
- **Estudos de Caso**



NÃO TEMOS DADOS

T

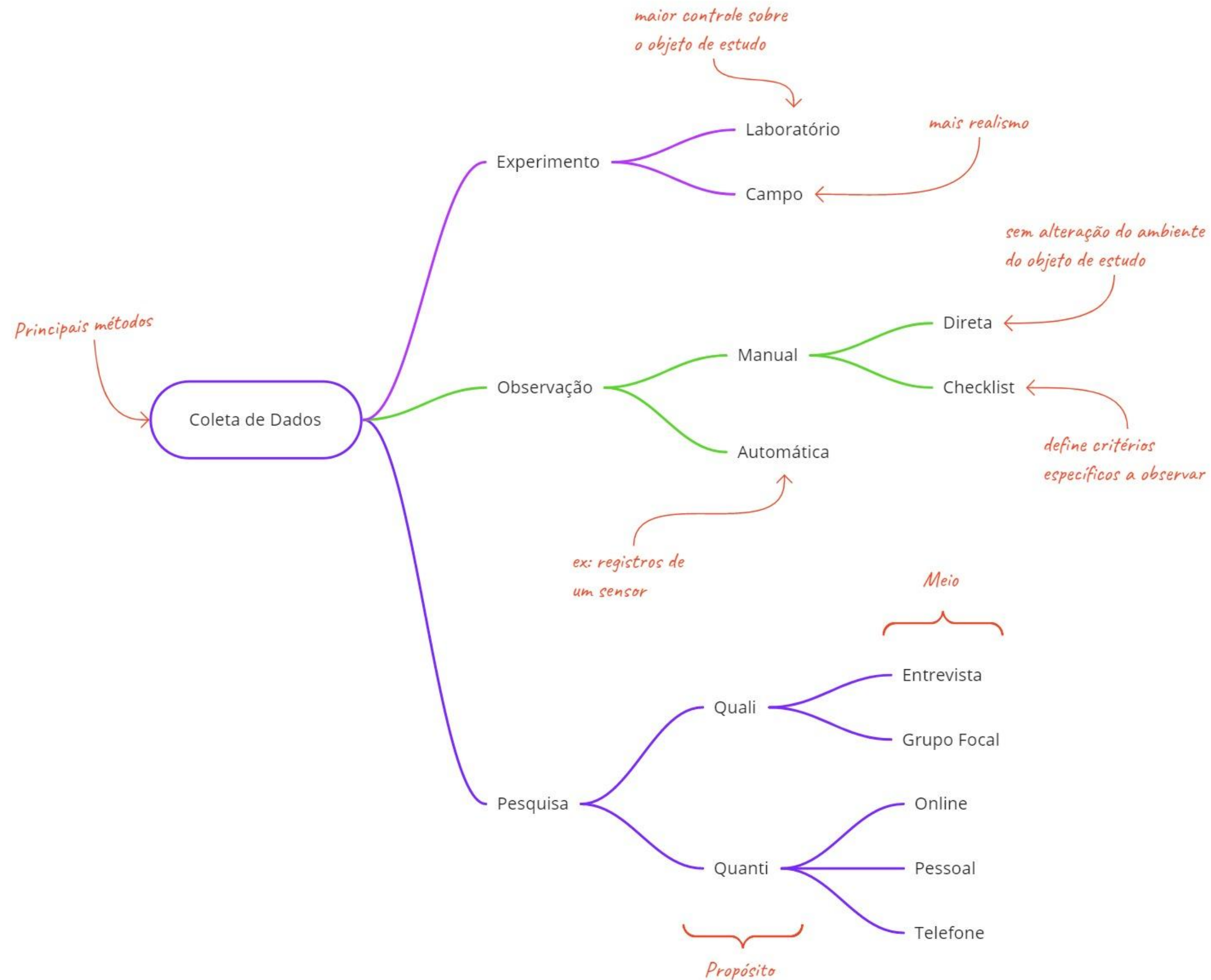
Dados não Disponíveis

O que fazer quando os dados não estão lá para suportar sua análise?

Atividade: Cite um exemplo de como endereçar esse problema.



T Coleta de Dados



T Coleta de Dados

1

Experimento

Lançar um novo produto em um cenário autêntico, para identificar o tamanho do mercado e como os consumidores e revendedores reagem ao manuseio, uso e recompra do produto.

2

Observação

Registrar todas as interações que os clientes fazem no e-commerce, desde a navegação até o check-out.

3

Pesquisa

Envio de um questionário com perguntas pré-estabelecidas para quantificar o interesse de potenciais clientes em um novo produto.



AGORA TEMOS DADOS

T

Dados Disponíveis

Os dados existem. E agora?

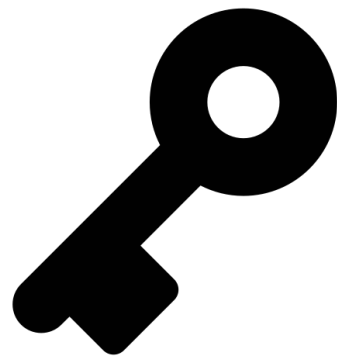
Atividade: Quero prever quem será o melhor jogador de futebol masculino do mundo. Posso usar os dados ao lado? Por que?



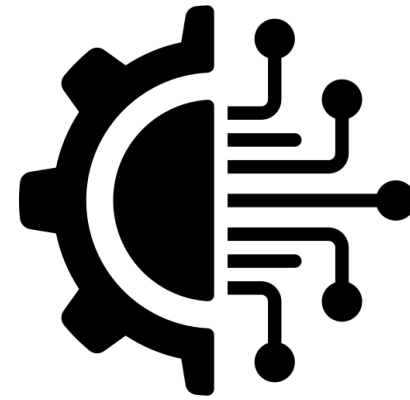
FIFA
ULTIMATE TEAM

T Dados Disponíveis

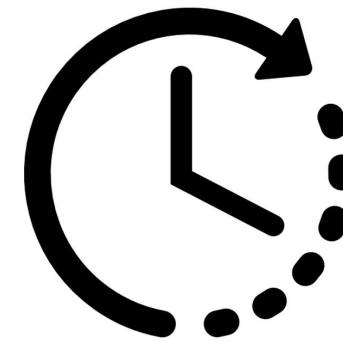
Contexto



Responsabilidade



Fonte



Atualização



Significado

T Dados Disponíveis

1- Como foi coletado?

- Qual foi o processo?
- Quem coletou?
- Qual o motivo? Será que tem algum viés do observador ou do processo?
- Os dados têm qualidade? Tem dado faltando ou dado incorreto?
- Sei o que significa cada dado?

T Dados Disponíveis

2- Qual a natureza dos dados?

- Qual o formato? Está estruturado? Qual a granularidade? É um dado agregado?
- É processável? Digital? Imagem ou video? HTML ou PDF?
- É armazenável? Preciso de um banco de dados ou uma planilha serve? É uma API?

T Dados Disponíveis

3- Os dados estão atualizados?

- Qual a frequência de atualização?
- Existe algum sistema de versionamento?
- Qual o processo de atualização?
- Existe algum monitoramento?

T Dados Disponíveis

4- E o acesso?

- Quem libera o acesso? Ou é livre de restrição?
- Preciso me identificar?
- O acesso é completo ou à apenas parte dos dados?
- Meu acesso expira?
- Posso compartilhar com terceiros?

T Dados Disponíveis

5- E quanto ao custo?

- O acesso é gratuito ou pago? Quanto custa?
- Caso seja pago, como é o modelo de precificação? Pagamento único, por tempo, por volume?
- Sendo gratuito, existem Direitos Autorais? Ou acordo de não divulgação?

T Dados Disponíveis

6- Atende as necessidades?

- Os dados precisam de enriquecimento?
- Tenho que fazer algum tratamento nos dados?
- Até quando ficam disponíveis?



O QUE FAZER COM OS DADOS?

Interfaces

Absolutamente tudo o que acontece dentro e fora da sua empresa gera um dado. Quando uma venda é feita, cada dado pode ser importante para ser usado pela gestão de logística.

Da mesma forma, quando um cliente entra no seu site, dados como a localização, quanto tempo ficou na página e quais produtos visitou também são informações valiosas se soubermos como usá-las.



Banco de Dados

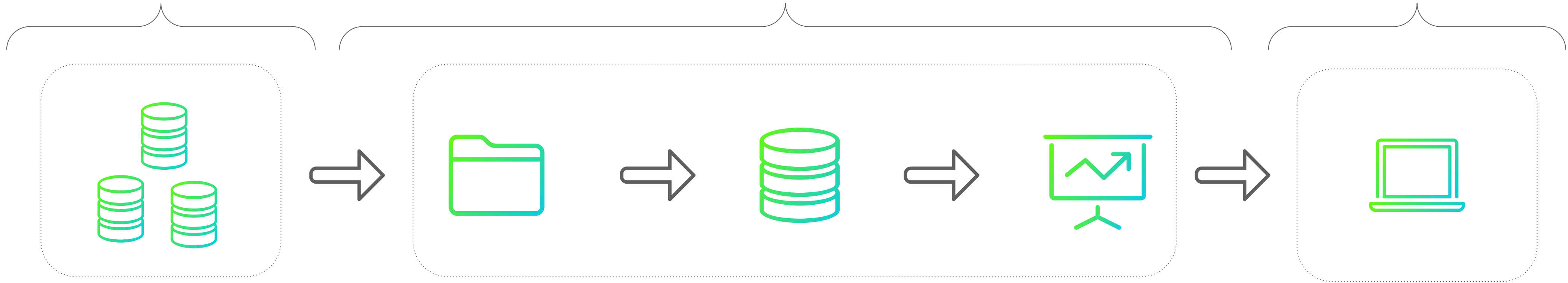
Em um banco de dados relacional, cada linha na tabela é um registro com uma ID exclusiva chamada chave. As colunas da tabela contém atributos dos dados e cada registro geralmente tem um valor para cada atributo



Instância Produtiva

Instância Analítica

Instância de Trabalho





Data Lake

É um tipo de repositório que armazena conjuntos grandes e variados de dados brutos em formato nativo. Com os data lakes, você tem uma visão não refinada dos dados. Essa estratégia de gerenciamento é cada vez mais usada por empresas que querem um grande repositório holístico para armazenar dados.

T


A vertical bar with a gradient from bright green at the top to cyan at the bottom, positioned to the left of the title.


ESTUDOS DE CASO

House of Cards





★★★★★ 2013 TV-MA 1 Season HD 5.1

Sharks gliding ominously beneath the surface of the water? They're a lot less menacing than this Congressman.

 This winner of three Emmys, including Outstanding Directing for David Fincher, stars Kevin Spacey and Robin Wright.



Because you watched Orange Is the New Black



Because you watched Red Lights

