



CIÊNCIA DE DADOS

CIÊNCIA OU CONJUNTO DE TRUQUES?

PEDRO H. DE BRITTO BATISTA - PROGRAMAÇÃO PARA ANÁLISE DE DADOS

PERGUNTA INICIAL

Se aplicarmos técnicas estatísticas e aprendizado de máquina em um conjunto de dados, estamos fazendo ciência ou apenas um conjunto sofisticado de truques?



PRINCIPAIS TÓPICOS

01

Evolução histórica da Ciência de Dados

Como a Ciência de Dados surgiu? Desde as primeiras discussões sobre análise de dados até seu reconhecimento como um campo independente, destacando marcos históricos essenciais.

02

Diferentes perspectivas (estatística, computacional e humana)

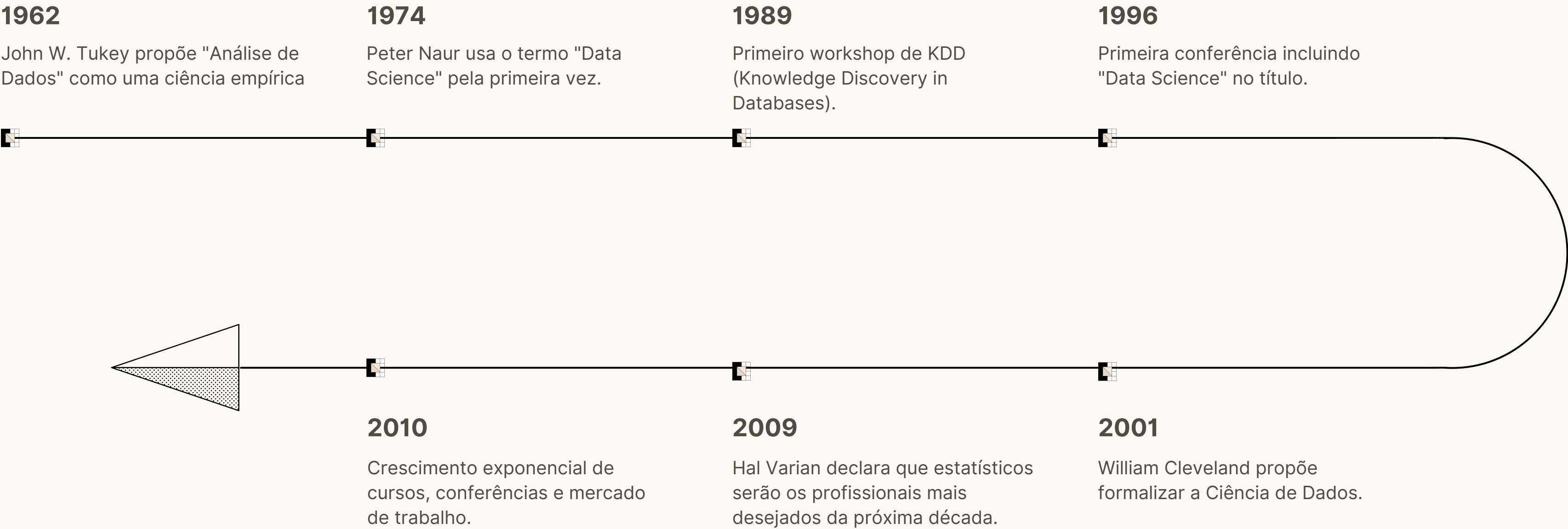
Como a Ciência de Dados é vista sob três óticas principais: a estatística, que fornece a base matemática para a modelagem; a computacional, que permite escalabilidade e eficiência; e a humana, que envolve interpretação, comunicação e ética.

03

Reflexão sobre o futuro da área

A Ciência de Dados continuará evoluindo para se consolidar como um campo científico estruturado? ou se será apenas uma fase da evolução tecnológica, além dos desafios futuros como transparência, ética e fundamentação teórica.

EVOLUÇÃO HISTÓRICA



TRÊS PERSPECTIVAS

Estática



Computação



Fator Humano





ESTÁTISTICA

A Favor

- A Ciência de Dados amplia o escopo da estatística, aplicando-a a problemas modernos e combinando-a com novas abordagens computacionais.
- Métodos estatísticos são fundamentais para garantir que a análise de dados seja rigorosa e confiável.

Contra

- A maior parte das técnicas de Data Science já existia dentro da estatística.
- Muitas abordagens de aprendizado de máquina são aplicadas de maneira empírica, sem um embasamento estatístico sólido.



COMPUTAÇÃO

A Favor

- Sem os avanços computacionais, o processamento de grandes volumes de dados seria inviável.
- Modelos complexos como redes neurais profundas só se tornaram práticos com o aumento do poder de processamento.

Contra

- A computação moderna facilita a aplicação de métodos, mas não define a Ciência de Dados como uma disciplina científica independente.
- Muitos modelos são usados sem plena compreensão de seus fundamentos teóricos, o que pode levar a erros.



FATOR HUMANO

A Favor ✓

- A interpretação dos resultados e a comunicação com stakeholders são essenciais para transformar análise de dados em decisões concretas.
- Ética e transparência são fundamentais para evitar vieses e garantir que os modelos sejam justos e explicáveis.

Contra ✗

- Se um modelo é preciso, mas não pode ser explicado, ele pode ser questionado do ponto de vista científico.
- A subjetividade na interpretação dos resultados pode levar a distorções ou decisões equivocadas.

CONCLUSÃO

- A Ciência de Dados integra estatística, computação e expertise de domínio.
- Ainda não tem uma estrutura teórica própria como outras ciências.
- O futuro pode consolidá-la como ciência ou mostrar que é apenas uma fase da evolução tecnológica.

PERGUNTAS PARA REFLEXÃO

- Ciência de Dados precisa de mais fundamentação teórica?
- O avanço computacional substitui o rigor estatístico?
- O que diferencia um cientista de dados de um analista de dados tradicional?

O B R I G A D O !