

# Bigdata, machine learning e outros animais: oportunidades e desafios para a defesa da concorrência

Paulo Furquim de Azevedo

INSPER

# Análise estatística convencional

- Dados estruturados
  - Estrutura pode variar (e.g. série de tempo, painel)
- Como regra geral, quanto mais dados, melhor
- Porém frequentemente não há dados suficientemente bons
  - ex

# Bigdata e outros: definição

## Bigdata

*a buzzword*

## Qual é a marca distintiva?

Não é “muito dado”

Via de regra: dados não estruturados

(textos, imagens, registros automáticos)

# Blockchain

- Blockchain: procedimento de registro de informações, verificáveis e compartilhadas
- Muitas das atividades tem sido registradas por meio de blockchain (e.g. bitcoin, comércio eletrônico).
- Muitas oportunidades de uso de dados

# Problema e a solução (ML)

- Problema: dados não-estruturados não são tratáveis pelos métodos tradicionais
  - Hiperdimensionalidade
  - Machine learning e estatística bayesiana para reduzir dimensionalidade
  - Ex: sistemas de classificações

# Um exemplo: notas fiscais eletrônicas

- Convênio entre Insper e SEFAZ-SP
- O que potencialmente pode ser feito?
  - Identificação de fraude em licitações
  - Screening de carteis
  - Estimação de demanda em atos de concentração
- Dificuldades e obstáculos