

# O uso de Big Data em Defesa da Concorrência



Carlos Ragazzo

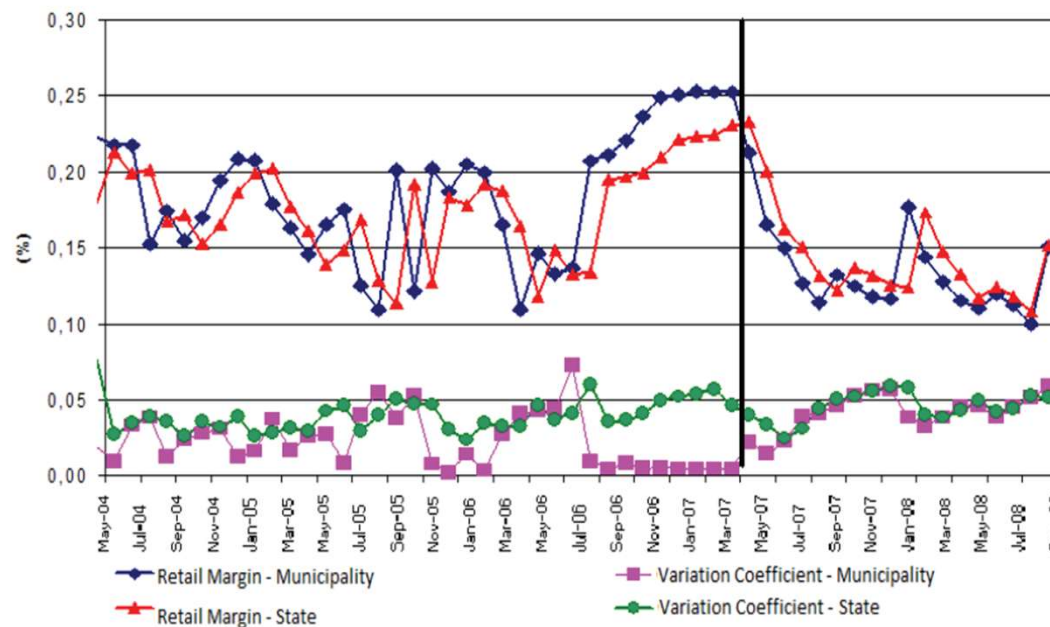


## Histórico da Aplicação de Filtros



# Brasil

## Combustíveis



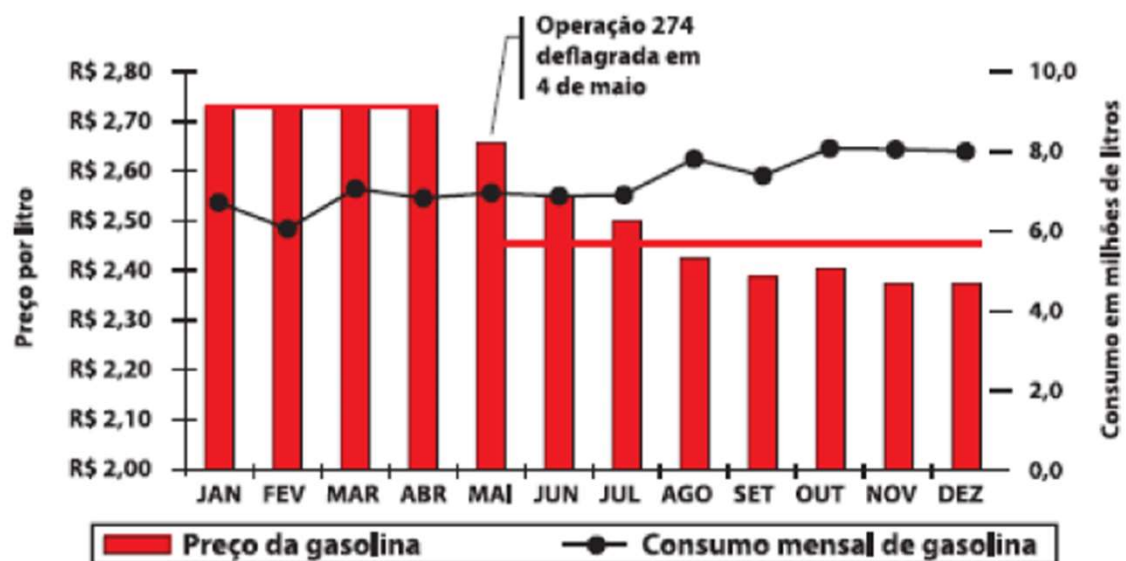
- Durante o cartel:
  - A margem de venda é maior que a referência regional;
  - Variância de preços ínfima.
- Após a descoberta do cartel:
  - A margem de venda é menor que a referência regional;
  - Variância de preços significativa.



Operação 274

## Investigações - início


Evolução do preço e do consumo, de gasolina tipo C, em João Pessoa





No "Café com o Presidente", Lula ressaltou que "é importante as pessoas ficarem atentas". "A gasolina não aumenta nada no posto. Se algum posto estiver aumentando, as pessoas podem denunciar porque não aumenta nada, e o óleo diesel aumenta 8,8%".

**A declaração resultou na elevação do número de denúncias**



Uso de bases de datos para diseñar  
políticas públicas



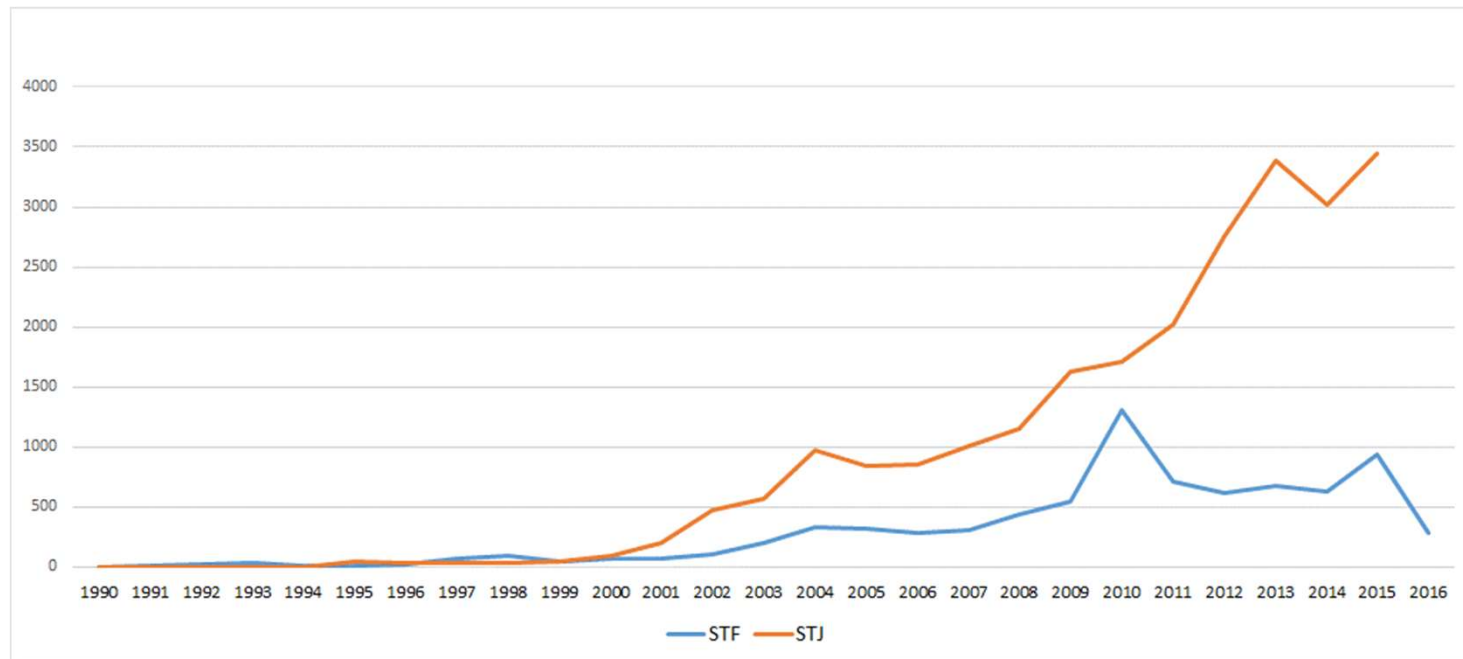
O projeto surgiu como iniciativa para aliar habilidades jurídicas e informáticas para produzir dados inéditos sobre o Supremo Tribunal Federal

Principais focos: dados sobre **andamentos dos processos**, sua **duração**, seus **atores**, suas **origens geográficas**, seu **assunto** e as **regularidades e correlações entre esses e outros elementos**.



## Exemplo de pesquisa:

**Evolução da quantidade de processos de improbidade no STF e no STJ (1990-2016)**



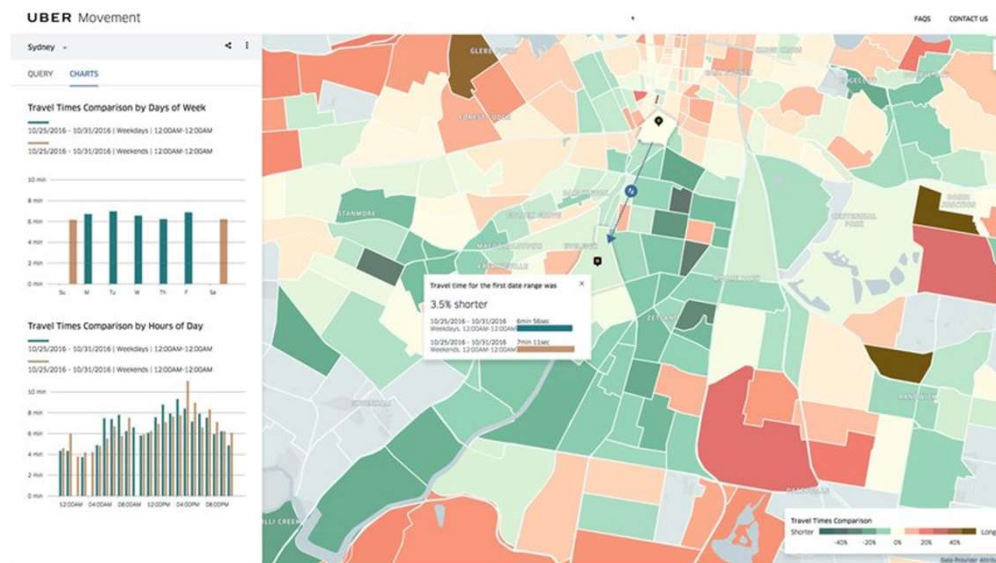
twitter  poderá ajudar no combate à dengue




Pesquisadores da UFMG desenvolveram um **software capaz de relacionar os tweets de pessoas informando que estão com dengue com os dados oficiais do governo.** Assim, poderá ser possível acompanhar em tempo real a disseminação da doença e focar o combate ao mosquito nas áreas mais atingidas.



# Uso de dados disponibilizados pelo aplicativo para mapeamento e criação de políticas públicas para melhorar o trânsito



Uber Movement



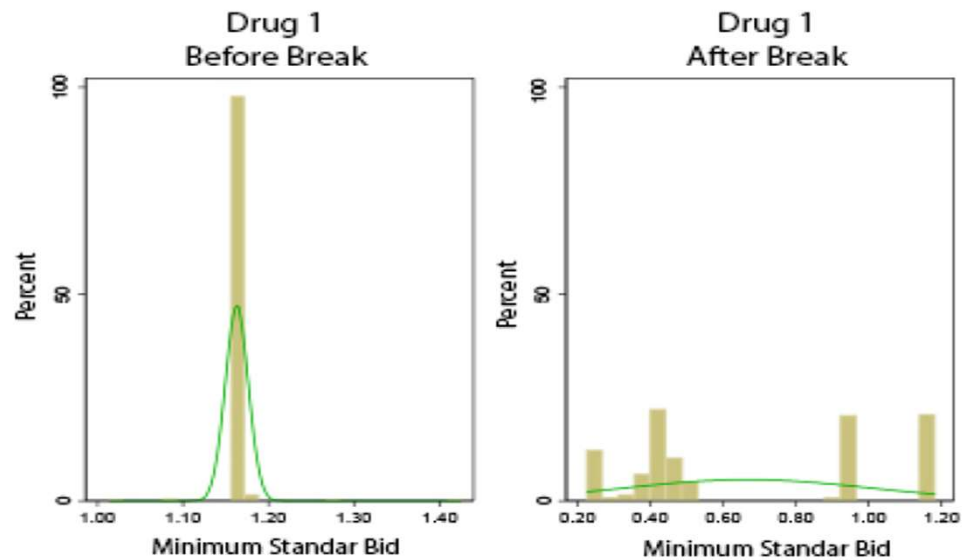
Uso de bases de dados para  
concorrência



México

Medicamento  
S

Durante o cartel:



- **Divisão do mercado estável;**
- **Variância de lances baixa.**



Projeto Cérebro





## Etapas do Projeto

Benchmarking  
da experiência  
internacional

Criação de  
bases  
metodológicas  
para uso de  
filtros  
econômicos

Coleta e  
estruturação  
de repositório  
de dados

Desenvolvimento  
e teste de  
algoritmos

Base de  
dados:

Baixar e integrar bancos de dados disponíveis:



etc.



# Criação da base de dados

Dump Oracle Comprasnet  
200GB

CSV ANP Georef  
2GB

TXT Petronect  
90GB

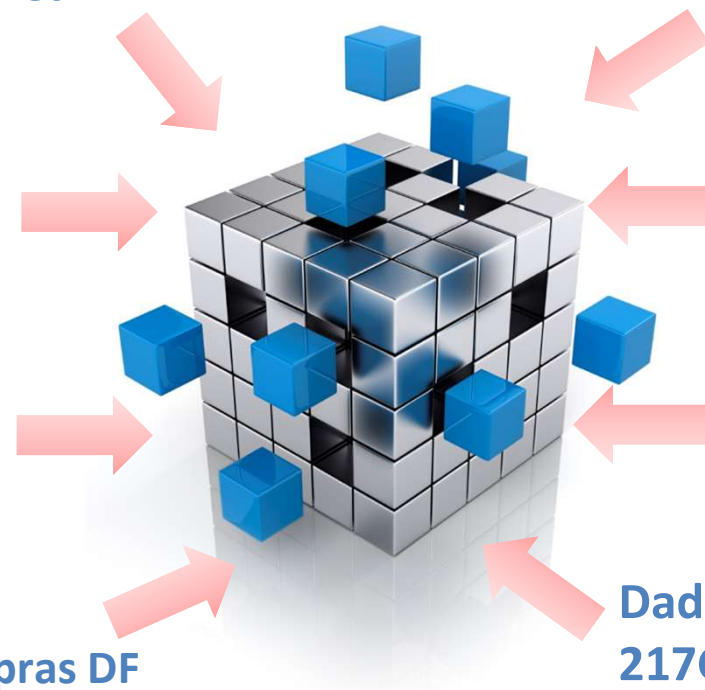
WebScraping Compras DF  
1GB

XLS CEMED ANVISA  
0,2GB

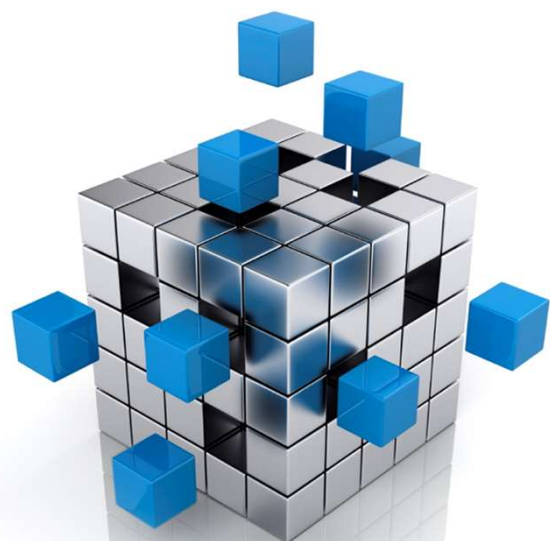
CSV Tribunal SP  
170GB

BAK Tribunal PE  
1,5GB

Dados Receita Federal  
217GB



## Criação da base de dados



>2TB  
~40 bases

### Licitações

- Gov. Federal, DF, RJ, RS, PB, PI, CE, MG, Petronect, Bco do Brasil

### Execução

- PE, SP, DF, RJ, RS, PB, ES, GO, PI, CE, MG

### Outras

- RAIS, ANVISA, CMED, ANPGeo, Receita

# Como utilizar dados?

Mineração dos  
dados

Testes  
estatísticos





## Como desenvolver algoritmos?

**Mineração dos dados**

- Fatos suspeitos / raros
- Sinais de competição simulada
- Similaridades e padrões nos comportamentos de concorrentes

**Testes estatísticos**

- Filtros econômicos: problemas de classificação/teste de hipóteses

## Fatores a serem investigados:



Propostas idênticas



Baixa competitividade



Análise de vitoriosos



Índice de fracasso



IPs idênticos



Rotatividade



Padrão de lances  
modificado



Coincidências cadastrais



Pregoeiro e homologador



Ganha mas não leva



Grandes quantias



Clusterização

## Fatores que podem apontar indícios de comunicação entre rivais:



Propostas idênticas



IPs idênticos



Coincidências cadastrais



Falta de aleatoriedade

## Fatores que podem apontar indícios de cobertura:



Índice de fracasso



Ganha mas não leva



Grandes quantias

## Fatores que podem apontar eventos fora do esperado:



Padrão de lances modificado



Rotatividade



Baixa competitividade





## Desafios no uso de Big Data:

- Manutenção da atualidade e Integração das bases;
- Mapeamento permanente dos pontos cegos;
- Evolução de funcionalidades.

# Obrigado!

[carlos.ragazzo@fgv.br](mailto:carlos.ragazzo@fgv.br)

[www.carlosragazzo.com.br](http://www.carlosragazzo.com.br)



Carlos Ragazzo

 **FGV DIREITO RIO**  
ESCOLA DE  
DIREITO DO  
RIO DE JANEIRO