

(21/05/2018) Entrevista para a Comunidade de Compras Públicas da Enap com o pesquisador da Fundação Getúlio Vargas Rafael Braem Velasco sobre *Data Analytics* em Compras Públicas. As perguntas foram elaboradas pelo professor da Enap Thiago Bergmann.

1)Recentemente, sua pesquisa foi premiada no CNMP. Pode nos contar sobre o que foi pesquisado e quais os resultados encontrados?

Velasco: A pesquisa busca compreender como técnicas de processamento e análise de grande volume de dados podem ser utilizadas para mensurar e comparar a qualidade das compras governamentais em nível municipal, estadual e federal. A pesquisa também busca compreender como que modelos de análise de dados podem ser utilizados para identificar padrões indicativos de irregularidades no uso de recursos públicos. Nosso principal objetivo é a construção de indicadores compostos calculados no nível de análise fornecedor e prefeitura. Deste modo, já calculamos, por exemplo, para cerca de 30.000 empresas fornecedoras de prefeituras de alguns Estados brasileiros, as seguintes informações: (i) valor médio de majoração contratual via aditivos; (ii) frequência média de contratação com base nas diferentes modalidades licitatórias da Lei 8.666; (iii) identificação de fornecedores que possuem alta probabilidade de ser “empresa de fachada”; e (iv) taxa de vitórias em licitações.

2)Quais as principais dificuldades que você encontra para trabalhar com dados de compras públicas de prefeituras e governos estaduais?

Velasco: Existe hoje no Brasil um cenário de ausência de sistematização na alimentação dos bancos de dados de Tribunais de Contas nos quais estão armazenados os dados sobre compras públicas de órgãos municipais e estaduais. Logo, existem grandes discrepâncias na qualidade dos dados encontrados em diferentes estados.

Gostaria de enfatizar dois problemas específicos em relação aos dados de compras públicas municipais: (i) falta de padronização; e (ii) dados incompletos.

Com relação à falta de padronização, ao compararmos os dados de compras públicas mantidos por Tribunais de Contas de diferentes estados percebemos que determinadas camadas de dados prioritárias para análises de irregularidades em licitações não existem nos bancos de dados de alguns órgãos. Exemplificativamente, o sistema de um TCE brasileiro que avaliamos não requer que o gestor municipal registre os dados a respeito de cada empresa que participou das licitações. Logo, em relação a este estado, ficam muito limitadas as possibilidades de análises de conclusão em compras públicas para todas as licitações de forma automatizada. Por exemplo, seria inviável identificar automaticamente se algum licitante neste Estado é uma empresa de fachada que possivelmente foi utilizada somente para dar aparência de competitividade ao certame licitatório.

Este exemplo aponta para o fato de que muitos Tribunais de Contas podem estar deixando de coletar dados prioritários para análises de qualidade do gasto público e identificação de indícios de corrupção. A coleta de vários outros dados relevantes também não é feita de forma padronizada ao redor do Brasil, tais como: regras dos editais, detalhes das propostas de cada licitante e detalhes a respeito dos aditivos posteriormente celebrados.

Com relação à incompletude dos dados percebemos que nas bases de alguns Tribunais de Contas existe um número significativo de prefeituras que não alimentaram o sistema do TCE com ao menos um contrato em um respectivo ano. Logo, parece haver indícios de que as unidades gestoras municipais não tratam com o devido nível de rigor a alimentação de dados de compras públicas.

Porém, estes casos de total ausência de dados de contratos refletem apenas a manifestação mais óbvia do problema de incompletude das bases dos TCE. É necessário identificar, também, se o conjunto de contratos adicionados por uma Prefeitura em um determinado ano se encontra dentro do valor que seria esperado como “completo”. Afinal, cada prefeitura, de acordo com o tamanho de sua população e o nível de desenvolvimento socioeconômico do município, irá apresentar um padrão esperado de gastos com contratações públicas. Os desvios dos valores esperados podem ser creditados à incompletude da base de dados do TCE, em razão da omissão do gestor municipal responsável por alimentar o sistema.

Logo, estamos desenvolvendo metodologias que analisem os valores gastos com contratações públicas em prefeituras de modo a identificar os desvios das distribuições esperadas. Em outras palavras, precisamos ser capazes de identificar aqueles municípios que alimentam a base do TCE com um conjunto de contratos em valor significativamente inferior ao esperado. Deste modo, será possível compreender com detalhes o real cenário de incompletude das bases de dados de compras públicas dos Tribunais de Contas do país. Vale ressaltar que com a utilização de dados de Notas Fiscais Eletrônicas é possível fazer uma validação cruzada entre a base de compras públicas do TCE e a base de NFe da Receita, de modo a aferir se existem discrepâncias entre os valores de notas emitidas por fornecedores da prefeitura e os valores de contratos inseridos na base do TCE.

Felizmente, uma vez que nossas metodologias de validação dos dados de TCE e as rotinas de limpeza da base estejam testadas e validadas, o custo marginal de analisar os dados de um novo Estado será baixo.

3) A análise de dados como conhecemos hoje só foi possível com o avanço do poder computacional. Quais software você conhece ou utiliza para proceder suas análises de dados ?

Velasco: Procuramos desenvolver soluções utilizando apenas soluções *open-source*. Atualmente desenvolvemos praticamente todos os modelos em Python e nossos bancos de dados relacionais estão em PostgreSQL. Alguns integrantes da equipe tinham maior familiaridade com R ou Matlab, mas consideramos que o Python vem atendendo muito bem a todos as nossas necessidades.

4) Na sua visão, como a análise de dados pode aprimorar as compras públicas e melhorar a qualidade nos gastos?

Velasco: Primeiramente, a análise de grande volume de dados permite que a administração pública identifique diversas oportunidades de economia de recursos. Exemplificativamente, estamos trabalhando com dados de compras de medicamento por prefeituras do Estado da Paraíba e constatamos grandes discrepâncias nos valores pagos pelos mesmos produtos. Suponho que parcela significativa destas discrepâncias não seja fruto de irregularidades ou má-fé, mas sim da incompletude das informações disponíveis para o gestor público municipal no momento da tomada de decisão e uma legislação de licitações excessivamente formalista. Adicionalmente, também é possível identificar claras oportunidades de economia de recursos mediante a centralização da demanda e obtenção dos respectivos ganhos de escala, ao invés da realização de várias licitações fracionadas.

Em segundo plano, a análise de grande volume de dados permite uma mudança considerável no monitoramento do gasto público por órgãos de controle, facilitando consideravelmente a identificação de casos de potenciais irregularidades que demandam maiores investigações, permitindo, por exemplo:

(i) identificar de forma automatizada empresas fornecedoras de governos municipais ou estaduais que possuem sócios “laranjas”;

(ii) identificar de forma automatizada fornecedores de governos municipais ou estaduais que são “empresas de fachada”;

(iii) detectar conluio entre participantes de licitação;

(iv) detectar fraudes em folha de pagamento de prefeitura, tais como servidores já falecidos (cruzando com SISOB) e servidores com múltiplos vínculos incompatíveis com a realidade (funcionário fantasma); e

(v) identificar casos de utilização de recursos de Convênio entre Governo Federal e prefeitura em despesas estranhas ao seu objeto.

5) Quais são algumas boas práticas, dentro da administração pública brasileira, de uso de dados para melhora da qualidade do gasto público e fiscalização?

Velasco: Alguns exemplos particularmente bem-sucedidos vêm à mente. No âmbito do Poder Executivo, venho acompanhando o relevante trabalho da Central de Compras do Ministério do Planejamento, que conseguiu gerar consideráveis economias de recurso a partir de centralização e racionalização da demanda, o que requer uma modelagem econométrica de grande volume de dados.

No âmbito dos órgãos de controle, os trabalhos da CGU e TCU no uso de dados para monitoramento do gasto público são muito avançados, com a criação de sistemas de informação relevantes como o DGI Consultas, ALICE e o painel da matriz de risco da rede Infocontas. Adicionalmente, o Ministério Público da Paraíba, por meio da equipe do GAECO, em parceria com a Universidade Federal de Campina Grande, também vem se destacando como um dos principais exemplos de uso de dados para monitoramento do gasto público no Brasil, com resultados muito exitosos.

6) Para finalizar, quais os próximos desafios da análise de dados na compras públicas?

Velasco: Gostaria de focar em três desafios específicos.

O primeiro se refere à gestão de dados de compras públicas no Brasil. Parece-me fundamental que todos os dados de compras públicas dos 5.570 municípios e 27 estados sejam centralizados em um mesmo banco de dados, gerido, por exemplo, pelo Ministério do Planejamento. Estes dados possuem elevado valor estratégico para a gestão pública e precisam ser geridos de forma padronizada e integrada.

O segundo desafio se refere ao fato de que hoje diversas camadas de dados relevantíssimas para análises de qualidade do gasto público não são sistematicamente coletadas ao redor do Brasil, tais como: termos de referência, editais de licitação e versões integrais de contratos. Para a maior parte dos municípios estes dados estão em formato não estruturado – i.e. dados em papel em uma gaveta de alguma repartição pública.

Por fim, o terceiro desafio se refere à criação de sistemas de informação capazes de evitar que o gasto público de baixa qualidade ou a irregularidade ocorra antes mesmo de qualquer dano ao erário estar concretamente materializado. Para tanto, a criação de banco de dados nacional de compras públicas seria um importante primeiro passo, pois, a partir de então, será possível regulamentar com detalhes a forma como cada unidade gestora deverá alimentar este banco no momento da celebração de cada novo contrato. Afinal, precisamos de um sistema de informação único, com abrangência nacional, no qual todos os órgãos públicos brasileiros deverão adicionar, em tempo real, os dados de cada novo edital de licitação aberto e de cada novo contrato celebrado. Certamente um projeto com estas características é viável e possui elevado valor para o país. O exemplo do sistema ALICE mostra que tecnicamente é plenamente factível desenvolver uma solução deste tipo para todo o Brasil.

Rafael Velasco é graduado em Direito pela FGV Direito Rio e mestre em Instituições, Política e Governo pela FGV-EBAPE. Ele foi coordenador de pesquisas do Programa de Transparência Pública (PTP-FGV) entre 2014 e 2016. Sua dissertação de mestrado examina a existência de discriminação em respostas a pedidos de acesso à informação baseada na identidade do requerente. Seus interesses de pesquisa incluem a utilização de análise de dados para a prevenção de corrupção, *open data* e políticas de transparência pública.

Thiago Bergmann de Queiroz possui graduação em Matemática e em Ciências Contábeis, além de mestrado em Administração pela Universidade de Brasília (UnB). Atualmente é analista judiciário do Tribunal Superior Eleitoral. Tem experiência na área de Administração, com ênfase em Ciências Contábeis. É professor e palestrante da Enap desde 2013.