

## Quanto Custa: Banco de Preços de Compras Públicas do Estado de Minas Gerais

Lucas G. L. Costa<sup>1</sup>, Marco Túlio Dutra<sup>1,2</sup>, Gabriel P. Oliveira<sup>1</sup>,  
Mariana O. Silva<sup>1</sup>, Daniela Cruz Soares<sup>3</sup>, Luciana de Cassia Silva Faria<sup>3</sup>,  
Wagner Meira Jr.<sup>1</sup>, Gisele L. Pappa<sup>1</sup>

<sup>1</sup>Universidade Federal de Minas Gerais (UFMG) – Belo Horizonte, MG – Brasil

<sup>2</sup>Universidade Federal de Ouro Preto (UFOP) – Ouro Preto, MG – Brasil

<sup>3</sup>Ministério Público do Estado de Minas Gerais (MPMG) – Belo Horizonte, MG – Brasil

lucas-lage@ufmg.br, marco.dutra@aluno.ufop.edu.br  
{gabrielpoliveira,mariana.santos}@dcc.ufmg.br  
{daniela,lfaria}@mpmg.mp.br, {meira,glpappa}@dcc.ufmg.br

**Abstract.** *This paper introduces Quanto Custa, a system developed by the Prosecution Service of the State of Minas Gerais in partnership with the Federal University of Minas Gerais to facilitate the querying and analysis of prices for public procurement items in the state. The system integrates data from bids, contracts, and invoices, and uses external sources such as ANVISA, ANP, and CEASA-MG to establish reference prices. Through features including item disambiguation and price variation analysis, Quanto Custa aims to enhance transparency and combat fraud in public procurement.*

**Resumo.** *Este artigo apresenta a ferramenta Quanto Custa, desenvolvida pelo Ministério Público do Estado de Minas Gerais em parceria com a Universidade Federal de Minas Gerais, para facilitar a consulta e análise de preços de itens de compras públicas no estado. O sistema integra dados de licitações, contratos e notas fiscais, além de utilizar fontes externas como ANVISA, ANP e CEASA-MG para estabelecer preços de referência. Através de funcionalidades como desambiguação de itens e análise de variação de preço, o Quanto Custa visa aumentar a transparência e combater fraudes nas compras públicas.*

### 1. Introdução

O volume de compras públicas no Brasil tem apresentado um aumento constante ao longo dos anos. Especificamente no estado de Minas Gerais, desde 2014, o número de licitações públicas vem crescendo continuamente, alcançando mais de 430 mil licitações até 2022.<sup>1</sup> Este vasto volume de dados torna a consulta de itens comprados e a investigação de superfaturamentos processos consideravelmente custosos e desafiadores. Para solucionar esses desafios, o Ministério Público do Estado de Minas Gerais, através do Gabinete de Segurança e Inteligência (GSI/MPMG), em parceria com o Departamento de Ciência da Computação da Universidade Federal de Minas Gerais (DCC/UFMG), desenvolveu o sistema Quanto Custa, no âmbito do Programa Capacidades Analíticas (PCA).

<sup>1</sup><https://fiscalizandocomtce.tce.mg.gov.br/>



**Figura 1. Tela inicial da ferramenta Quanto Custa.**

O Quanto Custa é uma ferramenta Web que disponibiliza mais de 30 milhões de itens de compras públicas para consulta, abrangendo licitações, contratos e notas fiscais. Além disso, o sistema inclui preços de referência obtidos de fontes externas como a Agência Nacional de Vigilância Sanitária (ANVISA), a Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) e Centrais de Abastecimento de Minas Gerais (CEASA-MG). O sistema não se limita apenas à consulta de preços, mas também oferece estatísticas gerais e gráficos temporais que auxiliam o monitoramento de preços. Uma funcionalidade adicional importante é a detecção de sobrepreço em compras públicas, que possibilita a identificação de possíveis alertas de fraudes.

## 2. Quanto Custa

O Quanto Custa é uma ferramenta prática e eficiente para a consulta de preços de itens de compras públicas em Minas Gerais. Com o objetivo principal de oferecer informações atualizadas sobre produtos e serviços adquiridos pelo estado e pelos municípios, a ferramenta também desempenha um papel fundamental na identificação de compras públicas que apresentam sobrepreço, uma prática fraudulenta em licitações, auxiliando na auditoria desses processos e no combate à corrupção no setor público. A Figura 1 apresenta a tela inicial da ferramenta, onde estão presentes os campos de busca e a opção para acessar a tela de filtros. A ferramenta está disponível apenas para os usuários internos do MPMG, mas há um vídeo institucional de apresentação do sistema.<sup>2</sup>

### 2.1. Dados

Para garantir a precisão e a atualização das informações, o Quanto Custa utiliza fontes de dados confiáveis. O sistema utiliza dois tipos de fontes de dados:

- **Compras públicas.** São considerados licitações municipais e estaduais, notas fiscais e contratos municipais. Os itens de licitações, notas fiscais e contratos municipais são obtidos do Sistema de Informações sobre Contratos Públicos (SICOM) do Tribunal de Contas do Estado de Minas Gerais (TCE-MG). Já os itens de licitações estaduais são

<sup>2</sup>Vídeo institucional do Quanto Custa: <https://youtu.be/Hm8KHYi-2sc>

obtidos por meio do Portal da Transparência do TCE-MG. Como o órgão que disponibiliza os dados de compras públicas é o TCE-MG, o Quanto Custa apresenta dados apenas dos municípios e do estado de Minas Gerais.

- **Preços de referência** No total, três fontes externas são consideradas para fornecer preços de referência, incluindo ANVISA (preços de medicamentos), ANP (preços de combustíveis) e CEASA-MG (preços de hortifrúti).

**Curadoria de dados.** Os dados coletados frequentemente apresentam problemas como valores anômalos, dados faltantes, falta de padronização e dados mascarados. Para assegurar a qualidade e a confiabilidade das informações no Quanto Custa, foi necessário implementar um rigoroso processo de curadoria de dados. Especificamente, foram implementadas as validações de qualidade propostas por [Oliveira et al. 2022, Oliveira et al. 2023]. Tais validações envolvem a identificação de registros problemáticos e a aplicação de métodos para corrigir ou, se necessário, remover esses dados do sistema.

**Desambiguação de itens.** Como a descrição do item de uma compra pública é um campo de texto livre, cada município preenche os dados de forma específica, sem um padrão. Isso causa a existência de itens iguais descritos de maneiras diferentes, dificultando a comparação de produtos e serviços similares. Portanto, para resolver esse problema, um arcabouço de desambiguação de itens foi implementado. O arcabouço utiliza técnicas de processamento de linguagem natural (PLN) e heurísticas para padronizar e estruturar as descrições dos itens [Silva et al. 2023, Brum et al. 2024]. A combinação dessas abordagens permite ao sistema Quanto Custa realizar comparações eficazes entre itens de compras públicas, melhorando a detecção de sobrepreço e a análise de mercado.

## 2.2. Arquitetura da Ferramenta

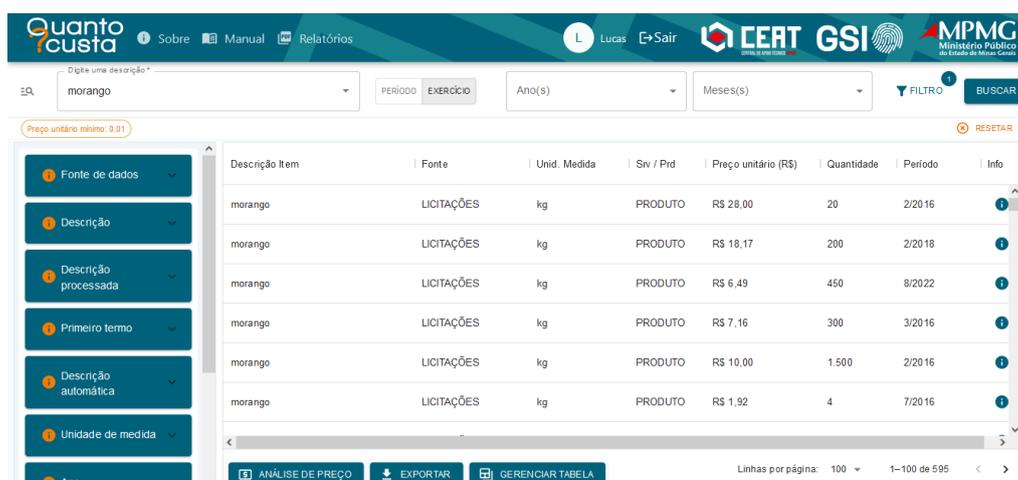
A ferramenta foi desenvolvida em quatro camadas, utilizando um conjunto de tecnologias adequadas para atender aos requisitos e objetivos estabelecidos. As tecnologias foram selecionadas com base em sua adequação aos requisitos do sistema Quanto Custa. As camadas e as tecnologias utilizadas em cada uma delas são descritas a seguir.

**Front-end.** Implementado no framework React, utilizando as linguagens de programação e marcação (JavaScript, HTML, CSS), que possibilita desenvolver interfaces interativas e responsivas. O WSO2 IS é utilizado para autenticação, permitindo uma autenticação segura por meio de um sistema de login.

**Back-end.** Implementado através de uma API em Python, com o auxílio do framework FastAPI, que proporciona rapidez e eficiência no desenvolvimento de APIs RESTful. A plataforma WSO2 API Manager é utilizada para gerenciar a segurança da API do sistema, garantindo a proteção e controle de acesso adequados.

**Armazenamento.** O armazenamento e gerenciamento eficiente dos dados são fornecidos pelo Elasticsearch, utilizado como banco de dados. Este oferece recursos avançados de busca e indexação, permitindo consultas rápidas e flexíveis.

**Mineração de Dados.** Scripts em Python são utilizados para implementar os arcabouços de desambiguação de itens e a curadoria de dados. Além disso, o Apache Spark é empregado para auxiliar na manipulação de big data e processamento distribuído, enquanto o Apache Hive serve como banco de dados para armazenar a saída gerada pelos arcabouços.



Descrição Item	Fonte	Unid. Medida	Srv / Prd	Preço unitário (R\$)	Quantidade	Período	Info
morango	LICITAÇÕES	kg	PRODUTO	R\$ 28,00	20	2/2016	1
morango	LICITAÇÕES	kg	PRODUTO	R\$ 18,17	200	2/2018	1
morango	LICITAÇÕES	kg	PRODUTO	R\$ 6,49	450	8/2022	1
morango	LICITAÇÕES	kg	PRODUTO	R\$ 7,16	300	3/2016	1
morango	LICITAÇÕES	kg	PRODUTO	R\$ 10,00	1.500	2/2016	1
morango	LICITAÇÕES	kg	PRODUTO	R\$ 1,92	4	7/2016	1

Figura 2. Tela com os resultados de uma busca de itens de compras públicas.

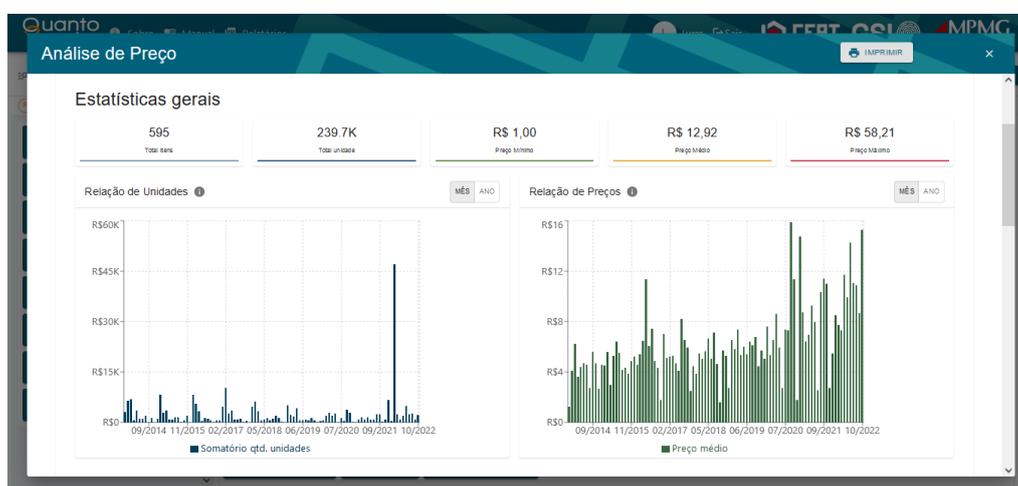


Figura 3. Tela com a análise de preço dos itens buscados.

### 2.3. Principais Funcionalidades

O sistema Quanto Custa foi desenvolvido com diversas funcionalidades que visam facilitar a consulta e análise de itens de compras públicas. A seguir, são apresentadas as principais funcionalidades implementadas.

**Busca de Itens.** Essa funcionalidade principal permite a aplicação de filtros simples como ano e mês, e avançados, como preço unitário, tipo de licitação, região, entre outros. Além dos filtros, duas formas de consultar itens de licitação foram implementadas: busca inteligente e busca assistida. A busca inteligente realiza a correspondência de texto com no mínimo 70% de precisão, enquanto a busca assistida permite escolher vários tipos de restrições entre os termos buscados (“Todas as palavras”, “Pelo menos uma das palavras”, “Não contendo as palavras” e “Com a frase exata”), além de permitir aplicar operadores lógicos para uma busca mais complexa. A Figura 2 apresenta um exemplo de um resultado de uma busca realizada no Quanto Custa.

**Análise de Preço.** Após o usuário realizar a consulta de um item licitado e aplicar os filtros desejados, a ferramenta permite que seja realizada uma análise de preço a partir

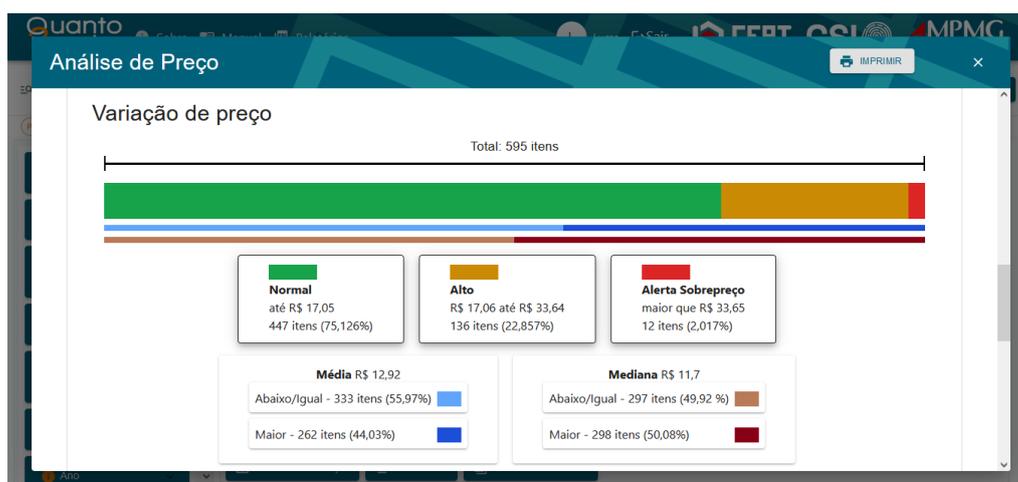


Figura 4. Tela com a variação de preço dos itens buscados.

dos itens retornados. Como indicado pela Figura 3, a tela de análise de preço exibe as seguintes informações: (i) *Estatísticas gerais*: total de itens, total de unidades, preço mínimo, médio e máximo, e (ii) *Visualizações temporais*: visualização com o histórico de preço e quantidade dos itens buscados.

**Variação de Preço.** A funcionalidade de variação de preço fornece, de forma visual, a variação de preço dos itens agregados na consulta. Além de indicar a quantidade de itens que estão com preços acima e abaixo da média e mediana, essa funcionalidade permite aos usuários identificar os itens com alerta de sobrepreço para que esses casos possam ser investigados pelos auditores. Essa funcionalidade é implementada utilizando a abordagem estatística proposta em [Silva et al. 2023, Silva et al. 2024]. Tal abordagem é baseada na amplitude interquartil para identificar itens com preços acima do esperado. A Figura 4 apresenta a variação de preço dos itens presentes na busca exemplificada pela Figura 2.

### 3. Estudos de Caso

Nesta seção, são apresentados dois estudos de caso com aplicações reais do Quanto Custa em tarefas cotidianas: consulta de preço de referência e detecção de sobrepreço.

**Consulta de Preço de Referência.** Neste cenário, uma pessoa usuária precisa obter um preço de referência para abertura de uma licitação de contratação de um palco cujas medidas são 6m x 4m. Assim, utiliza-se a funcionalidade de busca assistida, onde é possível aplicar restrições em relação aos termos buscados e operadores lógicos. Formula-se uma consulta que deve conter os termos “palco” e “6\*x\*4”, sem conter o termo “equipamento”. Dessa forma, a consulta é especificada para retornar apenas resultados relacionados a palcos, excluindo equipamentos de iluminação. O caractere \* indica que podem haver vários caracteres entre os números 6 e 4 e a letra x, contemplando todas as formas de escrever 6m x 4m (e.g., “6mx4m”, “6 x 4”, “6x4”). Aplica-se também um filtro para o campo “ano” com valores “2022” e “2023”, para que sejam retornados apenas itens de compras recentes. Após acessar a tela de análise de preço, é possível verificar um preço médio para esse item nas compras públicas já realizadas. Dessa forma, tal valor pode ser utilizado como preço de referência para elaborar uma nova licitação.

Descrição Item	Unid. Med.	Preço unitário (R\$)	Quantidade	Info
gasolina combustivel comum aba francisco sgasolina s	litro	<u>R\$ 561,00</u>	30.000	<a href="#">i</a>
gasolina comum cota reservado	litro	<u>R\$ 23,50</u>	15	<a href="#">i</a>
gasolina comum	litro	<u>R\$ 8,33</u>	85.000	<a href="#">i</a>
gasolina comum c automotivo de acordo com portaria anp	litro	<u>R\$ 7,89</u>	194.000	<a href="#">i</a>

**Figura 5.** Tela com a variação de preço de itens referentes a gasolina comum. Os itens com alerta de sobrepreço são mostrados no topo e seus valores são sublinhados em vermelho.

**Deteção de Sobrepreço.** Nesta tarefa, é necessário identificar a ocorrência de sobrepreço (i.e., itens adquiridos com valor acima do padrão) em compras de gasolina comum. Utiliza-se novamente a funcionalidade de busca assistida e monta-se uma consulta que deve conter os termos “gasolina” e “comum”, sem conter o termo “aditivada”. Dessa forma, a busca retorna apenas itens de gasolina comum, excluindo itens de gasolina aditivada. Em seguida, seleciona-se a unidade de medida “litro” e acessa-se a funcionalidade de variação de preço. Nela, é exibida a variação de preço dos itens, indicando aqueles com preços normais, altos e com alerta de sobrepreço. No exemplo ilustrado pela Figura 5, dois itens apresentam alerta de sobrepreço: um com o valor de R\$ 23,50 por litro e outro com R\$ 561,00 por litro. Na coluna “Info”, há um botão que abre o detalhamento da licitação, permitindo visualizar informações sobre a licitação para que ela possa ser encontrada em outros sistemas ou no próprio portal da transparência do município.

#### 4. Conclusão

Este artigo apresentou o Quanto Custa, uma ferramenta desenvolvida para realizar diversas buscas avançadas sobre itens de compras públicas e auxiliar na investigação de fraudes envolvendo sobrepreço. De fato, o sistema visa ser uma importante ferramenta para os servidores do Ministério Público do Estado de Minas Gerais (MPMG), facilitando a consulta e análise de produtos e serviços adquiridos em processos de compras públicas. Ao oferecer dados precisos e atualizados, o sistema não apenas auxilia na otimização dos recursos e na eficiência operacional, mas também desempenha um papel crucial na identificação de irregularidades e no combate à corrupção. Dessa forma, o Quanto Custa fortalece a capacidade investigativa do MPMG e contribui significativamente para a transparência e a integridade no setor público.

**Agradecimentos.** Ao Ministério Público do Estado de Minas Gerais pelo apoio através do Programa Capacidades Analíticas. Ao CNPq, CAPES e FAPEMIG pelo apoio aos pesquisadores envolvidos.

## Referências

- Brum, P. P. V. et al. (2024). Unsupervised grouping of public procurement similar items: Which text representation should I use? In *LREC-COLING*, pages 17176–17185, Torino, Italia. ELRA and ICCL.
- Oliveira, G. P. et al. (2022). Ferramentas open-source de qualidade de dados para licitações públicas: Uma análise comparativa. In *SBBD*, pages 116–127, Porto Alegre, RS, Brasil. SBC.
- Oliveira, G. P. et al. (2023). Assessing data quality inconsistencies in brazilian governmental data. *Journal of Information and Data Management*, 14(1).
- Silva, M. O. et al. (2023). Análise de sobrepreço em itens de licitações públicas. In *WCGE*, pages 118–129, Porto Alegre, RS, Brasil. SBC.
- Silva, M. O. et al. (2024). Overpricing analysis in brazilian public bidding items. *Journal on Interactive Systems*, 15(1):130–142.