

A Dataset of Public Procurement Auction Results from Brazil’s Compras.gov.br Platform (2003–2023)

Lucas Vicentim Perasolo^{*} and Christina Johanna Liepold¹

¹TUM School of Management, Technical University of Munich

^{*}lucas.vicentim-perasolo@tum.de

Last update: September 26, 2025

Abstract

We aim to facilitate access to a public dataset of procurement auctions conducted via Brazil’s Compras.gov.br platform from 2003 to 2023. The dataset available on GitHub comprises 170,096 auction records, 1,365,244 item records, 9,304,097 participant records, and 740,581 commitment records. To support accessibility and usability, the release includes unofficial English-language documentation and illustrative code examples. By providing these resources, we aim to lower barriers to entry and foster research in diverse fields related to public procurement and procurement auctions research in general.

Procurement is an important economic activity that drives demand for goods and services across both the public and private sectors worldwide. According to World Bank estimates, public procurement accounts for approximately 12% of global GDP¹². Despite its economic significance, access to procurement data remains limited. However, in recent years, digitalization has increased transparency by adopting electronic procurement systems, particularly in public procurement. Countries such as Brazil (via **Compras.gov.br**), Ukraine (via **ProZorro**), Chile (via **ChileCompra**), and India (via the **Government eMarketplace**) exemplify this shift toward more transparent procurement processes.

¹OECD panel data [Accessed 5 May 2025]

²World Bank study [Accessed 5 May 2025]

We highlight one such initiative, **Compras.gov.br**, the official electronic procurement platform of the Brazilian federal government, developed and maintained by the Ministry of Management and Innovation in Public Services. **Compras.gov.br** serves as a centralized system for conducting public procurement through various modalities, most notably, electronic procurement auctions (so-called *"Pregão Eletrônico"* in Portuguese), which are the predominant governmental procurement method in Brazil. The platform provides end-to-end support for procurement processes, including tender publication, supplier registration, bid submission, auction conduction, and contract award. By making procurement opportunities, tender processes, and outcomes publicly accessible, the system promotes transparency, efficiency, and competition. Through the Compras.gov.br interface, users can access detailed information on procurement notices, auction outcomes, supplier participation, and prices paid.

To facilitate access to and highlight the availability of governmental procurement data, we first provide an overview of current and future usage of public procurement data by research areas. To this end, Table 1 showcases a range of interdisciplinary research topics that leverage data from Compras.gov.br or closely related sources, illustrating the dataset’s broad potential usages for future studies. In addition, Table 2 in the Appendix summarises the dataset made available by this paper, which comes down to four CSV files covering procurement auctions, participating bidders, auctioned items, and related financial commitments, totalling approximately 2.5 GB in size.

Data Collection: To the best of our knowledge, there are two primary methods for interacting with the Brazilian government’s public procurement infrastructure. The first is through its official Application Programming Interface (API),³ which, while promising, appears to be under active development and may experience occasional instability. The second method involves either manually downloading procurement records from the **Portal da Transparência**⁴ or using its associated API.⁵ We generated our dataset via web scraping using the manual download interface of the Portal da Transparência website.

Data Overview: We present our dataset in its original form, with minimal preprocessing. We excluded government purchases conducted under procurement modalities other than electronic procurement auctions. No entries were removed due to missing values, as doing so would require assumptions and justifications beyond the scope of this paper. Moreover, to

³Overview of API and documentation [Accessed 5 May 2025]

⁴Portal da Transparência Downloads [Accessed 5 May 2025]

⁵Portal da Transparência API Documentation [Accessed 5 May 2025]

Research Domain	Applications & Supporting Literature
Auction Mechanisms & Economic Efficiency	<ul style="list-style-type: none"> • Auction design and theoretical analysis [7] • Auction duration effects on prices [12] • Government discretion and procurement efficiency [3] • Competition effects in large-scale auctions [13] • Transaction cost analysis from payment delays [8]
Network Analysis & Market Structure	<ul style="list-style-type: none"> • Firm co-bidding network characterization [10] • Hierarchical structures in procurement markets [15] • Corruption network analysis [16] • Defense industrial base network structure [2]
Machine Learning & Predictive Analytics	<ul style="list-style-type: none"> • Government purchase classification using ML [11] • NLP applications for procurement data [9] • Bidder recommendation systems [6] • Price prediction models [5]
Fraud Detection & Collusion Analysis	<ul style="list-style-type: none"> • Spatial analysis for bid rigging detection [1] • Machine learning-based collusion detection [14] • Market manipulation and bid-rigging impact analysis [4]

Table 1: Current and potential usages of public procurement data organized by research areas, with supporting literature demonstrating each application’s feasibility and academic relevance.

support users who may be unfamiliar with Portuguese (PT), we translated the column names into English (EN) and provided unofficial descriptions based on the official data dictionary.⁶ These are summarized in Tables 3, 4, 5, and 6 in the Appendix.

Data Availability & Examples: The complete dataset is publicly available in the GitHub repository: *A Dataset of Public Procurement Auction Results from Brazil’s ComprasNet Platform (2003–2023)*. The repository also includes a **examples** directory with use cases to help users explore and analyse the data. Note that these examples are intended for guidance and have not been optimised for analytics purposes.

Licenses: The dataset is released under the same licenses as the source; namely, the **Open Database License** and the **Database Contents License**.

⁶Translations are based on the official data dictionary available at Data Dictionary in original language [Accessed 5 May 2025]

Appendix: This section describes the dataset’s features, including translated column names (unofficial), data types, and completeness levels (i.e., the percentage of missing values for each variable). We aim to help users understand the structure of the data extracted from the Brazilian governmental procurement systems. In the tables below, SIASG stands for ”Sistema Integrado de Administração de Serviços Gerais” (”Integrated system for general services administration”), and CNPJ means ”Cadastro Nacional de Pessoas Jurídicas” (”National Registry of Legal Entities”).

CSV File	Number of Records	Number of Features	File Size
<i>Licitações</i> (Procurement Auctions)	170,096	17	74 MB
<i>Itens</i> (Auctioned Goods and Services)	1,365,244	14	297 MB
<i>Participantes</i> (Auction Participants)	9,304,097	13	1.9 GB
<i>Empenhos</i> (Related Commitments)	740,581	10	180 MB

Table 2: Summary of Datasets

Feature (PT)	Description (EN)	Completeness	Data Type
Número Licitação	Number that identifies the bidding in SIASG	100%	int64
Código UG	Code of the management unit responsible for the bidding	100%	int64
Nome UG	Name of the management unit	100%	object
Código Modalidade Compra	Code of the purchase modality	100%	int64
Modalidade Compra	Type of purchase modality	100%	object
Número Processo	Number of the bidding process	99.9%	object
Código Empenho	Code of the bidding commitment	100%	object
Data Emissão Empenho	Date of issuance of the commitment	88.1%	object
Observação Empenho	Observation/note of the commitment	100%	object
Valor Empenho (R\$)	Value of the commitment in Brazilian Reais	100%	object

Table 3: Detailed Info: *Empenhos* (Related Commitments)

Feature (PT)	Description (EN)	Completeness	Data Type
Número Licitação	Number that identifies the bidding in SIASG	100%	int64
Código UG	Code of the management unit responsible for the bidding	100%	int64
Nome UG	Name of the management unit	100%	object
Código Modalidade Compra	Code of the purchase modality	100%	int64
Modalidade Compra	Type of purchase modality	100%	object
Número Processo	Number of the bidding process	100%	object
Código Órgão	Code of the agency responsible for the bidding	100%	int64
Nome Órgão	Name of the agency	100%	object
Código Item Compra	Code of the item in SIASG (complex structured code)	98.95%	object
Descrição	Description of the item in SIASG	100%	object
Quantidade Item	Quantity of the item	100%	int64
Valor Item	Total value of the item	100%	object
Código Vencedor	CNPJ of the winning bidder	100%	object
Nome Vencedor	Name of the winning CNPJ	100%	object

Table 4: Detailed Info: *Itens* (Auctioned Goods and Services)

Feature (PT)	Description (EN)	Completeness	Data Type
Número Licitação	Number that identifies the bidding in SIASG	100%	int64
Código UG	Code of the management unit responsible for the bidding	100%	int64
Nome UG	Name of the management unit	100%	object
Código Modalidade Compra	Code of the purchase modality	100%	int64
Modalidade Compra	Type of purchase modality	100%	object
Número Processo	Number of the bidding process	100%	object
Objeto	What is being purchased, sold, or leased	100%	object
Situação Licitação	Current status of the bidding process	100%	object
Código Órgão Superior	Code of the superior agency responsible for the bidding	100%	int64
Nome Órgão Superior	Name of the superior agency	100%	object
Código Órgão	Code of the agency responsible for the bidding	100%	int64
Nome Órgão	Name of the agency	100%	object
UF	Brazilian state abbreviation	100%	object
Município	City where the bidding takes place	100%	object
Data Resultado Compra	Date of publication of the homologation	100%	object
Data Abertura	Date for submitting proposals	99.98%	object
Valor Licitação	Total bid value	100%	object

Table 5: Detailed Info: *Licitações* (Procurement Auctions)

Feature (PT)	Description (EN)	Completeness	Data Type
Número Licitação	Number that identifies the bidding in SIASG	100%	int64
Código UG	Code of the management unit responsible for the bidding	100%	int64
Nome UG	Name of the management unit	100%	object
Código Modalidade Compra	Code of the purchase modality	100%	int64
Modalidade Compra	Type of purchase modality	100%	object
Número Processo	Number of the bidding process	100%	object
Código Órgão	Code of the agency responsible for the bidding	100%	int64
Nome Órgão	Name of the agency	100%	object
Código Item Compra	Code of the item in SIASG	100%	object
Descrição Item Compra	Description of the item in SIASG	100%	object
Código Participante	CNPJ of the bidding participant	100%	object
Nome Participante	Name of the participant's CNPJ	100%	object
Flag Vencedor	Indicates if the participant is the winner ("SIM" / yes or "NÃO" / no)	100%	object

Table 6: Detailed Info: *Participantes* (Auction Participants)

References:

- [1] Ricardo Carvalho de Andrade Lima and Guilherme Mendes Resende. "Using the Moran's I to detect bid rigging in Brazilian procurement auctions". In: *The Annals of Regional Science* 66.2 (2021), pp. 237–254.
- [2] Thiago Caliari et al. "Defense industrial base and network structure: measurement, analysis and public policies for the Brazilian case". In: *Defence and Peace Economics* 36.3 (2025), pp. 373–394.
- [3] Dimas Fazio. "The two sides of discretion: Assessing efficiency and quality in government purchases". In: *Available at SSRN 4212969* (2022).
- [4] Dimas Fazio and Alminas Zaldokas. "Kamikazes in Public Procurements: Bid-Rigging and Real Non-Market Outcomes". In: *European Corporate Governance Institute–Finance Working Paper* 1054 (2025).
- [5] Manuel J Garcia Rodriguez et al. "Award price estimator for public procurement auctions using machine learning algorithms: case study with tenders from Spain". In: *Studies in informatics and control* (2021).
- [6] Manuel J García Rodríguez et al. "Bidders recommender for public procurement auctions using machine learning: Data analysis, algorithm, and case study with tenders from Spain". In: *Complexity* 2020.1 (2020), p. 8858258.
- [7] Paul Klemperer. "Auctions: theory and practice". In: (2018).
- [8] Matheus Pereira Libório et al. "Effect of late payments on the transaction costs of Brazilian public procurement carried out through electronic auctions". In: *International Journal of Procurement Management* 18.2 (2023), pp. 260–278.

- [9] Wesley Lima et al. “BidCorpus: A multifaceted learning dataset for public procurement”. In: *Data in Brief* 58 (2025), p. 111202.
- [10] Marcos S Lyra et al. “Characterization of the firm–firm public procurement co-bidding network from the State of Ceará (Brazil) municipalities”. In: *Applied Network Science* 6.1 (2021), p. 77.
- [11] Thiago Marzagão. “Using SVM to pre-classify government purchases”. In: *arXiv preprint arXiv:1601.02680* (2015).
- [12] Alexandre Oliveira, Abdoulaye Fabregas Masllovet, and Mihály Fazekas. “Auction Length and Prices: Evidence from Random Auction Closing in Brazil”. In: *World Bank Policy Research Working Paper* 8828 (2019).
- [13] Rodrigo de Oliveira Leite, Bernardo Spitz Paiva, and Luiz Claudio Sacramento. “Competition and efficiency in procurement auctions: evidence from a million Brazilian auctions”. In: *Journal of Public Finance and Public Choice* 40.1 (2025), pp. 2–22.
- [14] Manuel J García Rodríguez et al. “Collusion detection in public procurement auctions with machine learning algorithms”. In: *Automation in Construction* 133 (2022), p. 104047.
- [15] Niclas Frederic Sturm et al. “High earnings through firm influence: the role of hierarchical structures in public procurement”. In: *EPJ Data Science* 14.1 (2025), p. 27.
- [16] Harald Waxenecker and Christina Prell. “Corruption dynamics in public procurement: A longitudinal network analysis of local construction contracts in Guatemala”. In: *Social Networks* 79 (2024), pp. 154–167.