(07-002) - Analyzing Open Data Quality and Trends in Public Procurement Colombian Platform amid the Digital Revolution

Gomez Cabrera, Adriana [1]; Cortes Rey, Juan Sebastian [1]; Sanchez Rivera, Omar Giovanny [1]; Felizzola, Heriberto [2]; Gomez, Camilo [3]

[1] Pontificia Universidad Javeriana, [2] Universidad de La Salle, [3] Universidad de los Andes

The digital revolution and the increasing availability of data have transformed management processes across businesses, governments, and society. Governments, aligning with international standards, have established digital platforms to communicate public procurement progress to citizens. Several public data platforms, including those in the construction sector, store vital information on the allocation of public resources. Colombia's notable accessibility to open data facilitates the development of data science projects. The open data platform offers valuable insights into construction projects at various lifecycle stages, benefiting stakeholders. Valuable trends and variable relationships can be gleaned from this platform. This document focuses on analyzing information from the open data platform, examining included variables and providing trend examples. The results indicate that project performance correlates with factors such as project size, procurement method, and planning, with notable correlations between time and cost deviations.

Keywords: Construction projects; digital revolution; public procurement; data science; open data

Analizando la Calidad de los Datos Abiertos en la Contratación Pública Plataforma Colombiana en medio de la Revolución Digital

La revolución digital y la creciente disponibilidad de datos han revolucionado los procesos de gestión para empresas, gobiernos y la sociedad. Los gobiernos, alineados con estándares internacionales, han establecido plataformas digitales para comunicar el progreso de la contratación pública a los ciudadanos. Varias plataformas de datos públicos almacenan información crucial que detalla la asignación de recursos públicos, y el sector de la construcción no es una excepción. Colombia se destaca en la accesibilidad a datos abiertos, lo que facilita el desarrollo de proyectos de ciencia de datos. La plataforma de datos abiertos proporciona información valiosa sobre proyectos de construcción en diferentes etapas del ciclo de vida, ofreciendo una oportunidad excepcional para las partes interesadas. Se pueden extraer tendencias valiosas y relaciones entre variables. Este documento se centra en analizar la información de la plataforma de datos abiertos. El estudio analizó las variables incluidas en los datos abiertos y proporcionó ejemplos de tendencias y análisis que se pueden desarrollar. Los resultados muestran que el rendimiento del proyecto está relacionado con el tamaño del proyecto, la forma de contratación y la planificación, entre otros, y también hay una correlación entre las desviaciones de tiempo y coste.

Palabras clave: Proyectos de construcción; revolución digital; contratación pública; ciencia de datos; datos abiertos

Correspondencia: Adriana Gómez Cabrera, adranagomez@javeriana.edu.co

Agradecimientos: Pontificia Universidad Javeriana

1. Introduction

Infrastructure plays a pivotal role in fostering economic and social development by enabling transportation and production (Céspedes et al., 2017; Vega et al., 2019). This infrastructure encompasses rural, urban, and interurban roads, along with other essential projects crucial for ensuring access to vital services such as healthcare, education, and commerce. Consequently, it significantly enhances the quality of life for citizens (Angulo et al., 2020; Ruiz et al., 2016). The importance of infrastructure projects leads governments around the world to prioritize the allocation of financial resources, both for the construction of new projects and for the improvement of existing ones (Navarro-Moreno et al., 2022). Due to their significance, these projects must be developed and delivered successfully, within the required budget, schedule, and scope (PMI, 2021).

However, time and cost deviations are commonly observed in infrastructure projects, especially those financed with public funds. This is particularly significant in public projects, where the construction sector receives a considerable allocation from government budgets. In fact, one-third of government expenditures are directed toward contracts for delivering goods, infrastructure, and services to citizens. Given the constraints of public resources, effective public procurement becomes essential for strategic governance and service delivery, particularly in infrastructure development. Well-designed public procurement systems play a crucial role in achieving project objectives and ensuring efficient resource allocation. (Open Contracting Partnership, 2022). Therefore, strengthening the public procurement system is crucial in the fight against corruption, as it accounts for more than one-third of the general budget. Specifically, the citizens of Latin America and the Caribbean demand more transparent and efficient states in the planning and execution of all their public policies (Zuleta, 2019).

Transparency in public procurement has undergone significant improvements in recent years. This progress is motivated by the imperative to enhance competitiveness, integrity, and efficiency in procurement processes. Many countries worldwide have established platforms that report on the execution of public spending using open data, enabling researchers to analyze trends from this information. The literature suggests that transparent publication of public spending can lead to increased participation from bidders, a higher prevalence of non-local winners, and improved quality of contract execution (Fazekas et al., 2016).

Open data constitutes a fundamental strategy to promote transparency, integrity, efficiency, and accountability in public spending management (Adam et al., 2020; Duguay et al., 2023). To leverage the potential of open contracting data, it is necessary to carry out analytical tasks to extract information about the performance and corruption risk in the procurement process (World Bank, 2022). This information is crucial for underpinning the formulation of public policies, decision-making in contracting procedures, and monitoring conducted by citizens, civil society organizations, and government entities (Misuraca & Viscusi, 2014). However, data analysis processes face various challenges, including technical obstacles and those specific to the contracting domain (Hellberg & Hedström, 2015). Managing the extensive volume of data, both structured and unstructured, related to contracting represents one of these significant technical challenges, requiring advanced skills for extraction, cleaning, processing, and analysis(Ansari et al., 2022; González-Mora et al., 2020). Additionally, beyond possessing the technical skills for data management and analysis, a deep understanding of the regulatory framework governing public procurement is essential. This knowledge is vital for guiding analytical tasks, ensuring that data is translated into meaningful, useful, and value-added

information, surpassing the mere generation of models and statistics lacking practical relevance(World Bank, 2022).

Colombia faces challenges regarding corruption, as evidenced by its ranking of 99th out of 180 countries with a score of 39/100 on the Global Corruption Index by Transparency International. (Transparency International, 2022). However, the national government has launched an ambitious initiative to digitize project-related information. Since 2011, the adoption of the public procurement system SECOP (Sistema Electrónico de Contratación Pública) has been pivotal in this endeavor. This platform promotes the digitization, centralization, and automated analysis of data thereby contributing to mitigating factors that negatively impact the public procurement (Colombia Compra Eficiente, 2024). The SECOP implementation has been a significant achievement in the modernization of government procurement processes in Colombia (Arias et al., 2018). It has established a new standard in transparency and efficiency in these operations. By providing transparent publication of contract-related documents and facilitating streamlined access to information, SECOP plays a crucial role in ensuring transparency and efficiency in procurement processes (Colombia Compra Eficiente, 2024).

Given the availability of public data and recognizing that the National Government of Colombia has been implementing a large investment plan for the development of road construction projects (Andrade, 2017), this research analyzes the available public data for these types of projects to provide results that may be useful in decision-making. The availability of digital information allows the identification of trends, opportunities, and challenges in the current landscape of Colombian infrastructure in the context of the digital revolution. In this paper, the Public Procurement System in Colombia will be presented, with a focus on the construction phase of infrastructure projects to reveal its structure, quantity, and quality of data. Additionally, examples of the types of analysis and visualizations that can be performed based on this data will be included.

2. Objectives

Identify the information available on the Colombian public data platform concerning public procurement processes for road construction projects.

Analyze the relationships among the variables cost and time deviation and the available information in Colombia's public procurement data platform for road projects.

3. Methodology

The research methodology included four stages: first, gathering information from the open data platform; second, understanding the available data; third, preparing data for analysis; and fourth, analyzing the data (See Figure 1). Each of these stages is described below.

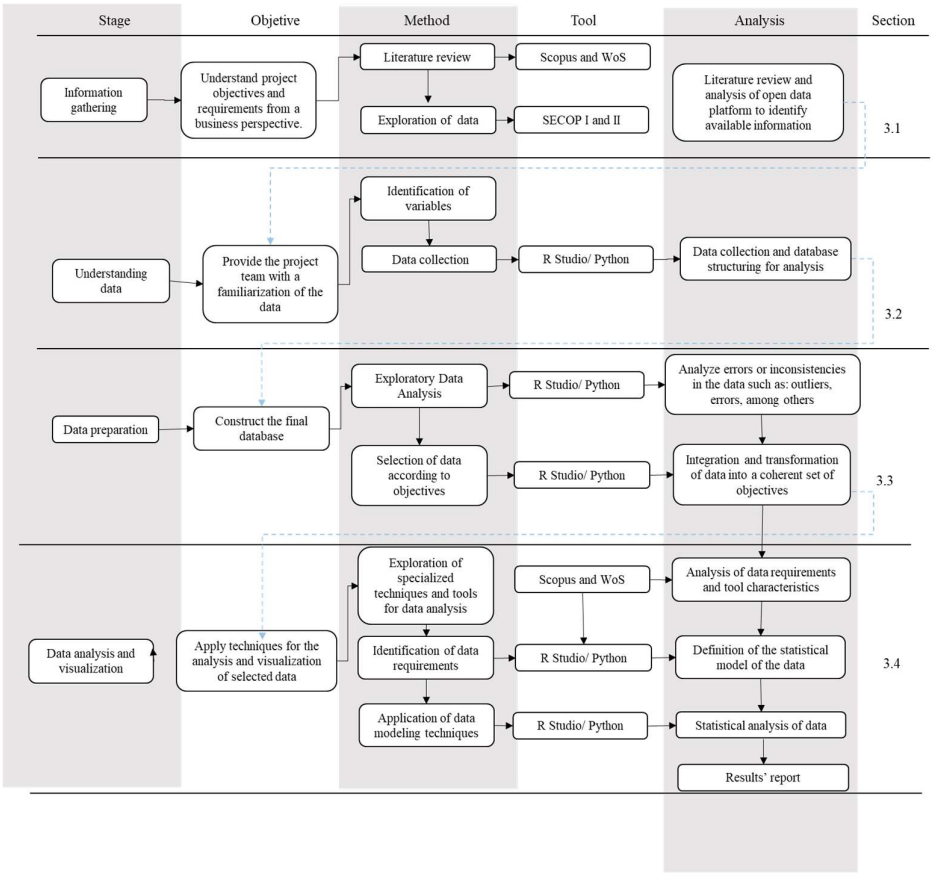

Figure 1: Research methodology.

3.1 Information gathering on the Open Data Platform.

Initially, a literature review was conducted on the variables analyzed in previous research related to time and cost deviations in infrastructure projects. Subsequently, an exploration of the SECOP open data platform was carried out to identify the available information. The consolidation of the filtered database was performed using algorithms to acquire information through programming, eliminating the need for manual search. The information related to road construction projects is included in two platforms: SECOP I, which serves as a platform for advertising process documents by contracting entities, and SECOP II, a new version designed as a transactional platform specifically for electronically managing all contracting processes. The required information was available on these platforms, including contractual data representing real data from public projects, supported by a regulatory framework mandatory for public procurement processes.

3.2 Understanding Data

In this stage, the process of understanding the information contained within the dataset was undertaken. This involved analyzing the structure, content of each column, and characteristics of the data to gain insights into its meaning and relevance for the research. Additionally, variables available in the SECOP that had also been included in previous investigations were identified.

3.3 Data Preparation

This stage involved cleaning, organizing, and structuring the data obtained from the SECOP platform to make it suitable for analysis. An exploratory data analysis was conducted to obtain information about the available variables and their nature. The objective was to understand the collected data, review whether the dataset's content ensures reliable data, or if it is necessary to correct some aspects, and identify the need to remove information that may bias the analysis (outliers). This stage was carried out following the methodology outlined by (Larose & Larose, 2015).

3.4 Data Analysis

During this stage, data visualization was developed, serving as a means of transforming complex data sets into clear and understandable representations. This facilitates their analysis, interpretation, exploration, and analytics (Camm et al., 2017; Munzner & Maguire, 2015). Dashboards integrate multiple sources of information offering a panoramic view through graphs, charts, and gauges that facilitate data interpretation (Few, 2013). In the context of public procurement, the goal of dashboards is to enable stakeholders to quickly and effectively monitor, analyze, and understand performance, key performance indicators, trends, and critical data related to procurement processes (World Bank, 2022). These tools help government agencies, policymakers, and citizens to detect patterns, trends, and anomalies, improving understanding of public spending, identifying corruption risks, and assessing the effectiveness of procurement practices(Agrahari & Srivastava, 2019; Barcellos et al., 2017).

In this stage, methods previously implemented were explored, aiming to find statistical learning tools useful for making sense of data (Gómez-Cabrera, 2021; Gómez-Cabrera et al., 2020; James et al., 2013). Bivariate analysis was conducted to establish the relationship between independent variables and dependent variables (time and cost deviations in this case).

Considering that the dependent variables are numerical, Spearman's rho (a non-parametric test) was calculated for the independent numerical variables after observing non-normality in the data. The null hypothesis, in this case, posits that there is no association between the two variables. (Alvarado & Obagi, 2008). For cases where the dependent variable is numerical and the independent variables are categorical, the Kruskal–Wallis test (a non-parametric test) was used after observing non-normality in the data. This test analyzes differences in the median values of groups. The null hypothesis in this case is that the population medians are equal for all groups (Gatignon, 2010). The Kruskal–Wallis test identifies if groups involved in the categorical variables exhibit different behavior concerning the dependent variables analyzed. However, it does not determine which of the groups the difference is statistically significant. Therefore, the Wilcoxon Mann-Whitney test (a non-parametric test) is used to compare paired data and establish this. The null hypothesis, in this case, is that the median difference between pairs of observations is zero, allowing the identification of categories with similar behavior that can be grouped (Goos & Meintrup, 2016). A p-value << 0.05 for all cases indicates that the null hypothesis must be rejected.

4. Results

This research included, first, the results of exploring the open data platform for road construction projects. Then, two case studies are conducted to identify factors contributing to time and cost deviations in secondary road construction projects in Colombia. These case studies include visualization of available data and bivariate analysis for secondary road construction projects.

4.1 Information gathering on the Open Data Platform.

Initially, an exploration of the data available on the Colombian state's public data platform was conducted. It was found that the data platform meets various international standards to promote transparency in disclosing information about government project contracts. One of these approaches is the "Open Contracting Data Standard," established by the "Open Contracting Partnership." Since its inception in 2015, the ODC Principles have been adopted by 96 national and sub-national governments (including Colombia) from around the world (Open Data Charter, 2022). Another measurement initiative, by the "Corruption Research Center Budapest," assesses data publication practices in 112 countries. Colombia is among the 82% of countries with a dedicated website for public procurement data, while 56% publish structured data on awarded contracts, and 32% offer downloadable data online. They identify 20 key characteristics for detecting corruption risk and other anomalies. Ukraine leads with all 20 characteristics available, followed by Colombia with 17(Fölsz & Tóth, 2018).

The data within SECOP is organized in a specific format, rendering them easily accessible, comparable, and suitable for analysis and monitoring. As structured data, their organization facilitates interoperability and reuse, thereby fostering greater transparency and efficiency in procurement processes in Colombia. Then, the data collection started searching information in the open data platform, www.datos.gov.co. To perform this task, a database containing various types of projects was downloaded, and those related to road construction were filtered. For each of these projects, the database included comprehensive information covering the entire project lifecycle. Figure 2 illustrates the variety of information that could be obtained and downloaded from the requested procurement process during the project lifecycle.

Figure 2: Available information in SECOP.

4.2 Data preparation

The search process involved filtering construction projects by criteria such as "type: construction," "family: roads," and "process state: settled" (projects that have completed their closing phase). For the first case study related to visualization, a comprehensive search of rural road construction projects was conducted, encompassing all project delivery methods, from 2014 to 2024 and analyzing original data available in SECOP.

For secondary roads, the data set includes projects from 2011 to 2022, encompassing 160 projects. This dataset also includes projects awarded through competitive processes and direct contracts. In this stage, unit value conversions of contracts were also conducted to consider

the value of money over time, converting this variable into legal minimum wages in Colombia. Additionally, the project's duration was standardized to days.

The independent variables, time and cost deviations, were included in this research following equations (Gransberg & Villarreal-Buitrago, 2002):

$$Time\ deviation = \frac{Final\ deadline - original\ deadline}{original\ deadline}$$

$$Cost\ deviation = \frac{Final\ cost - contract\ value}{contract\ value}$$

The independent variables for the bivariate analysis in this study are included in Table 1, indicating their description. The numerical variables included how they were measured and the different possible values for the categorical ones. These variables were considered in previous research (Gómez-Cabrera et al., 2020).

Table 1. Independent variables (Gómez-Cabrera et al., 2020)

Variable	Unit/ Values
Estimated project cost	Minimum salaries (*)
Project duration	Days
Project Intensity (How much money is spent per day (Federal Highway Administration, 2018)	Minimum salaries/days
Process Type (Modality chosen for project delivery)	Competitive Bidding, Abbreviated Selection, Minimum Contract
Geographical region	Amazonia, Andina, Caribe, Orinoquia, or Pacifica

(*) Estimated cost were converted to legal monthly minimum salaries in Colombia

4.3 Visual Analytics Tools for Rural Roads Contracts

This section discusses how data visualization can transform the monitoring and analysis of rural roads procurement in Colombia by leveraging the capabilities of a tool specifically designed for this purpose. By integrating and processing information from sources such as SECOP I and SECOP II, the tool not only simplifies the task of cleaning, standardizing, and processing procurement databases, but also transforms this data into actionable insights by generating procurement statistics and performance indicators. Through its user-friendly, intuitive, and easy-to-use interface, it facilitates the representation of data through graphs, maps and tables, allowing users a deep and detailed understanding of the dynamics and possible irregularities in rural roads procurement processes. The tool allows an analysis of contracting volume (in number and value of contracts) by date range, contracting regime, type of processes, type of contract, region, public entity, and supplier.

The first dashboard provides an overview of the overall figures on rural roads procurement processes, including the total number of contracts, the total value of procurements, the number of public buyers and the number of suppliers (see Figure 3). This information is crucial to evaluate the size of the contracting market and the volume of operations, as well as to appreciate the degree of participation and competition, considering the number of contractors at the national and regional levels. Additionally, the temporal evolution of contracting activities is presented, which allows identifying trends and patterns in rural road contracting over time.

This analysis is essential to understand the potential impact of government changes, both at the national and local levels, on contracting dynamics.

Figure 3: General summary of contracting processes.

Figure 4 illustrates the dashboard that facilitates the analysis of the contracting regime, the modalities used and the types of contracts. The purpose of this dashboard is to provide detailed information on the distribution and dynamics of interaction between the variables mentioned. For this purpose, a Sankey diagram is used, where the thickness of the lines represents the quantity or economic value of the contracts linked to each category in a proportional manner, thus enabling an immediate visual interpretation of the connections and identification of the predominant categories. The interactions revealed by the Sankey diagram may indicate systematic preferences for certain contracting modalities or types of contracts, which could reflect the level of competition and transparency of certain types of contracts and regimens for rural road contracting. For example, the graph shows a prevalence of non-competitive and in some cases restrictive processes such as direct contracting, minimum amount contracts and abbreviated selection, indicating the need to generate policies that encourage public entities to improve competition and transparency through open processes such as public bidding.

In Figure 5 the third dashboard shows a ranking of contractors and public entities based on statistics and indicators such as the number of contracts, the total value of contracts, the percentage of contracts with modifications and the percentage of contracts awarded through direct contracting. At the top left of the dashboard, two indicators show the total number of entities and contractors, according to the department filter. This visualization is intended to identify the entities and contractors of rural roads with the highest contracting volumes, both at the national and departmental levels. In addition, it allows identifying the entities with the highest percentage of contracts with modifications and those using non-competitive contracting mechanisms. This is significant information for citizen control and the activation of follow-up mechanisms by entity responsible for overseeing public expenditure in rural roads.

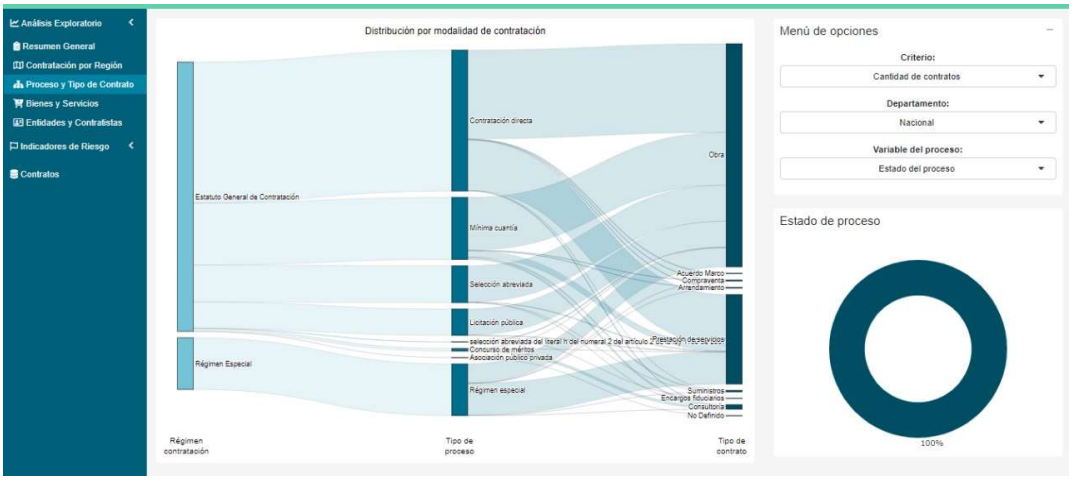

Figure 4: Public contracting by type legal regime, contract and modalities

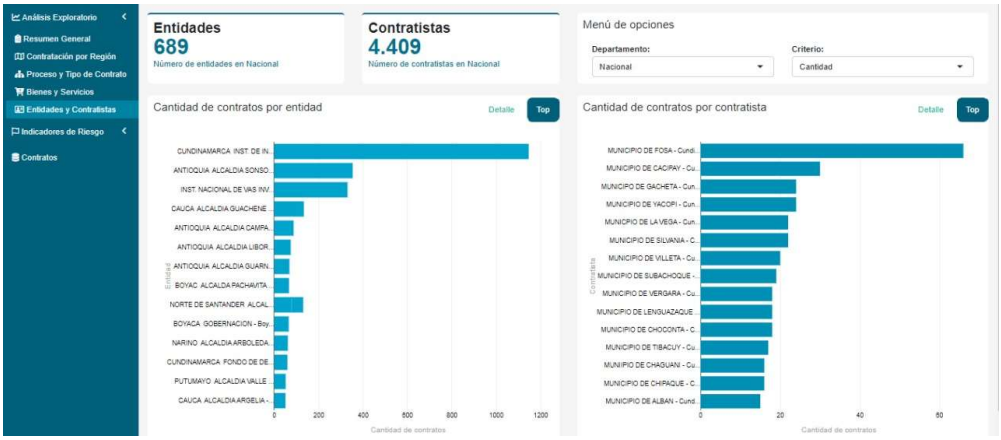

Figure 5: Public contracting by type legal regime, contract and modalities

4.4 Bivariate analysis

The bivariate analysis consists of comparing time and cost deviation with each independent variable presented in Table 1. Additionally, a comparison between time and cost deviations revealed a statistically significant positive relationship of 0.31.

For secondary roads, the significant relationships between the numerical variables for time deviation are included in **¡Error! No se encuentra el origen de la referencia.**, and for cost deviation in **¡Error! No se encuentra el origen de la referencia.**. In both cases, Spearman's rho is reported for instances where the p-value was less than 0.05. Project intensity is positively correlated in both cases. Additionally, the project duration exhibits a significant negative correlation, suggesting that greater values of contract duration result in higher time deviations. The estimated cost is significant for cost deviation suggesting that greater values of contracts result in higher cost deviations.

Table 2 Significant numerical relationships - Time deviation

Variable	Spearman Rho
Original duration	-0.31

Project intensity	0.30

Table 3 Significant numerical relationships - Cost deviation

Variable	Spearman Rho
Project intensity	0.29
Estimated cost	0.24

For secondary roads, the significant relationships between the categorical variables for time deviation are included in Table 4, and for cost deviation in Table 5. Following the identification of a significant relationship in each case, the Wilcoxon Mann-Whitney test was applied to analyze pairs of groups and determine which exhibit significantly different behavior. The process type is included in both cases, as the only significant relationship.

Table 4 Significant categorical relationships – Time deviation

Variable	Categories
Process type	Direct Contract
	Others (Abbreviated selection – Public bidding

Table 5 Significant categorical relationships – Cost deviation

Variable	Categories
Process type	Direct Contract
	Others (Abbreviated selection – Public bidding

5. Conclusions

The paper aimed to identify and explore the information in Colombian public data platforms, regarding road construction procurement, and analyze the behavior and relationships associated with cost and time deviations. Exploring Colombia's public procurement database revealed pertinent data on road construction projects meeting international standards. Colombia's OECD membership since 2018 reflects its commitment to top-notch public policies. Leveraging the information in this platform can enhance procurement transparency. However, raw data lacks depth for effective decision-making and informing society adequately. Thus, research of this kind provides valuable insights to advance this objective, benefiting stakeholders and society. Specifically, understanding the behavior and causes of time and cost deviations can inform the design of policies aimed at reducing inefficiencies in the provision of essential infrastructure.

The bivariate analysis for secondary road projects unveiled significant positive relationships for numerical variables like project intensity, affecting both time and cost deviation. Project intensity reflects the investment per unit of time and is determined during project planning. Consequently, inaccuracies in estimating project intensity may stem from shortcomings in this initial phase.

The bivariate analysis for secondary road projects unveiled significant relationships for categorical variables such as process type in both time and cost deviation. This underscores the impact of the chosen contractor award modality on project performance.

Ongoing research includes the application of inferential and machine learning models to expand analysis capabilities across several use-cases. Future research should involve other public data platforms to compare results across countries, including open data content and variable relationships.

6. References

Adam, I., Fazekas, M., & Tóth, B. (2020). *Measuring the benefits of open contracting - Case studies on Mexico, Paraguay, and Slovakia*. http://redflags.govtransparency.eu/wp-content/uploads/2020/01/GTI-WP_OC-benefits-research_final-report_20200121.pdf

Agrahari, A., & Srivastava, S. K. (2019). A data visualization tool to benchmark government tendering process. *Benchmarking: An International Journal*, *26*(3), 836–853. https://doi.org/10.1108/BIJ-06-2017-0148

Angulo, M., Bertelli, A., & Woodhouse, E. (2020). The political cost of public–private partnerships: Theory and evidence from Colombian infrastructure development. *Governance*, *33*(4), 771–788. https://doi.org/10.1111/gove.12443

Ansari, B., Barati, M., & Martin, E. G. (2022). Enhancing the usability and usefulness of open government data: A comprehensive review of the state of open government data visualization research. *Government Information Quarterly*, *39*(1), 101657. https://doi.org/10.1016/j.giq.2021.101657

Arias, C., Rengifo, D., Serrano, M., Perez, D., & Torrado, L. (2018). Analysis of the qualifying requirements for public procurement in road infrastructure projects terminated abnormally after being convened: The case of Valle del Cauca (Colombia)[Caso valle del cauca (Colombia)]. *Revista de Obras Publicas*, *165*(3598), 54–59.

Barcellos, R., Viterbo, J., Miranda, L., Bernardini, F., Maciel, C., & Trevisan, D. (2017). Transparency in practice. *Proceedings of the 18th Annual International Conference on Digital Government Research*, 139–148. https://doi.org/10.1145/3085228.3085294

Camm, J. D., Fry, M. J., & Shaffer, J. (2017). A Practitioner's Guide to Best Practices in Data Visualization. *Interfaces*, *47*(6), 473–488. https://doi.org/10.1287/inte.2017.0916

Céspedes, R., Rosero, J., Montaño, W., & Reyes, J. F. (2017). Methodology for defining the functionality of advanced measurement infraestructure in Colombia. *2017 IEEE PES Innovative Smart Grid Technologies Conference - Latin America, ISGT Latin America 2017*, *2017*, 1–6. https://doi.org/10.1109/ISGT-LA.2017.8126745

Colombia Compra Eficiente. (2024). *Colombia Compra Eficiente - Agencia Nacional de Contratación Pública* . Https://Www.Colombiacompra.Gov.Co/.

Duguay, R., Rauter, T., & Samuels, D. (2023). The Impact of Open Data on Public Procurement. *Journal of Accounting Research*, *61*(4), 1159–1224. https://doi.org/10.1111/1475-679X.12479

Fazekas, M., Tóth, I. J., & King, L. P. (2016). An Objective Corruption Risk Index Using Public Procurement Data. *European Journal on Criminal Policy and Research*, *22*(3), 369–397. https://doi.org/10.1007/s10610-016-9308-z

Few, S. (2013). *Information dashboard design: Displaying data for at-a-glance monitoring* (Analytics Press, Ed.; Second Edition).

Fölsz, H., & Tóth, I. J. (2018). *CORRUPTION_RESEARCH_CENTER_BUDAPEST*. https://www.crcb.eu/

Gatignon, H. (2010). Statistical Analysis of Management Data. In *Statistical Analysis of Management Data*. Springer New York. https://doi.org/10.1007/978-1-4419-1270-1_1

Gómez-Cabrera, A. (2021). *Identification of factors generating time and cost deviation in construction projects: a case study in rural roads in Colombia*. Universidad de los Andes .

Gómez-Cabrera, A., Sanz-Benlloch, A., Montalban-Domingo, L., Ponz-Tienda, J. L., & Pellicer, E. (2020). Identification of factors affecting the performance of rural road projects in Colombia. *Sustainability (Switzerland)*, *12*(18). https://doi.org/10.3390/SU12187377

González-Mora, C., Garrigós, I., & Zubcoff, J. (2020). *An APIfication Approach to Facilitate the Access and Reuse of Open Data* (pp. 512–518). https://doi.org/10.1007/978-3-030-50578-3_36

Goos, P., & Meintrup, D. (2016). *STATISTICS WITH JMP*. Wiley.

Gransberg, D. D., & Villarreal-Buitrago, M. (2002). Construction project performance metrics. *Annual Conference of AACE International*. https://www.proquest.com/docview/208181138?pq-origsite=gscholar&fromopenview=true

Hellberg, A.-S., & Hedström, K. (2015). The story of the sixth myth of open data and open government. *Transforming Government: People, Process and Policy*, *9*(1), 35–51. https://doi.org/10.1108/TG-04-2014-0013

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to statistical learning*. Springer. https://doi.org/10.1007/978-1-4614-7138-7_1

Larose, D. T., & Larose, C. D. (2015). Data Mining and Predictive Analytics (Wiley Series on Methods and Applications in Data Mining). *Wiley Series*, 794. https://doc.lagout.org/Others/Data Mining/Data Mining and Predictive Analytics %5BLarose %26 Larose 2015-03-16%5D.pdf

Misuraca, G., & Viscusi, G. (2014). Is Open Data Enough? *International Journal of Electronic Government Research*, *10*(1), 18–34. https://doi.org/10.4018/ijegr.2014010102

Munzner, T., & Maguire, E. (2015). *Visualization analysis and design* (Ebscohost, Ed.).

Open Contracting Partnsership. (2022). *Open Contracting Partnership*. https://www.open-contracting.org/

Open Data Charter. (2022). *Open Data Charter*. https://opendatacharter.net/who-we-are/

Ruiz, J. D. G., Arboleda, C. A., & Botero, S. (2016). A Proposal for Green Financing as a Mechanism to Increase Private Participation in Sustainable Water Infrastructure Systems: The Colombian Case. *Procedia Engineering*, *145*, 180–187. https://doi.org/10.1016/j.proeng.2016.04.058

Transparency International. (2022). *CORRUPTION PERCEPTIONS INDEX*. Https://Www.Transparency.Org/En/Cpi/2022.

Vega, L., Cantillo, V., & Arellana, J. (2019). Assessing the impact of major infrastructure projects on port choice decision: The Colombian case. *Transportation Research Part A: Policy and Practice*, *120*, 132–148. https://doi.org/10.1016/j.tra.2018.12.021

World Bank. (2022). *Using Data Analytics in Public Procurement Operational Options and a Guiding FrameworkEquitable Growth, Finance & Institutions Insight*. World Bank. https://doi.org/10.1596/37467

Zuleta, M. M. (2019). *Hacia una política de datos abiertos del Sistema de Compra Pública para los países miembros de la RICG*. https://ricg.org/wp-content/uploads/2019/11/Publicacion-hacia-una-pol%C3%ADtica-de-datos-abiertos-del-sistema-de-CP.pdf

Communication aligned with the Sustainable Development Goals