

ALGORITMO EN PYTHON PARA PREDECIR Y DETECTAR ANOMALÍAS EN  
PROCESOS CONTRACTUALES DE LA ALCALDÍA DE PUERTO RICO, META,  
MEDIANTE TÉCNICAS AVANZADAS DE MACHINE LEARNING

“Modalidad: Curso o seminario de profundización”



PAULA ALEJANDRA FONTECHA MURILLO



UNIVERSIDAD SANTO TOMÁS  
FACULTAD INGENIERIA INDUSTRIAL  
VILLAVICENCIO

2025

ALGORITMO EN PYTHON PARA PREDECIR Y DETECTAR ANOMALÍAS EN  
PROCESOS CONTRACTUALES DE LA ALCALDÍA DE PUERTO RICO, META,  
MEDIANTE TÉCNICAS AVANZADAS DE MACHINE LEARNING

“Modalidad: Curso o seminario de profundización”

PAULA ALEJANDRA FONTECHA MURILLO

Trabajo de grado presentado como requisito para optar al título de Ingeniero Industrial

Asesor

Mg. CRISTINA ISABEL COGOLLO TORRES

Magíster en Ingeniería Mecánica

UNIVERSIDAD SANTO TOMÁS  
FACULTAD INGENIERIA INDUSTRIAL  
VILLAVICENCIO

2025

**Autoridades Académicas**

**P. Álvaro José ARANGO RESTREPO, O.P.**

Rector General

**P. Mauricio Antonio CORTÉS GALLEGO, O.P.**

Vicerrector Académico General

**P. Luis Antonio ALFONSO VARGAS, O.P.**

Rector Seccional Villavicencio

**P. Adrián Mauricio GARCÍA PEÑARANDA, O.P.**

Vicerrector Académico Seccional Villavicencio

**Mg. Julieth Andrea SIERRA TOBÓN**

Secretaria General Seccional Villavicencio

**Mg. Víctor Andrés RINCÓN GONZALEZ**

Decano de la Facultad de Ingeniería Industrial

## **Agradecimientos**

Agradezco a la Universidad Santo Tomas de Villavicencio, por abrirme las puertas e irme formando poco a poco para lograr convertirme en el profesional que quiero ser, junto con los docentes y decano quienes han realizado un gran trabajo a lo largo de mi trayectoria, también quiero agradecer a mis compañeros, a mis padres por que fueron mi apoyo constante a la hora de perseguir mis sueños y este título será una muestra de esa fuerza y perseverancia que me inculcaron.

## Tabla de Contenido

Resumen.....	7
Abstract .....	8
Introducción .....	9
Título del proyecto .....	10
Planteamiento del caso.....	11
Marco de referencia .....	12
Preguntas de reflexión.....	15
Objetivos .....	16
Objetivo General .....	16
Objetivos específicos .....	16
Marco Metodológico.....	17
Tipo de estudio.....	17
Narración del caso.....	20
Lecciones y recomendaciones.....	35
Referencias bibliográficas.....	36

## Lista de figuras

<b>Figura 1</b> Metodología del algoritmo predictivo y de detección de anomalías en procesos contractuales de Puerto Rico, Meta. ....	19
<b>Figura 2</b> Cronograma de actividades metodológicas para el proyecto de predicción y detección de anomalías contractuales. ....	20
<b>Figura 3</b> Configuración inicial del entorno de trabajo en Python.....	21
<b>Figura 4</b> Carga del archivo de procesos contractuales en Google Colaboratory .....	21
<b>Figura 5</b> Definición del mapeo de columnas para la base de datos .....	22
<b>Figura 6</b> Recolección y limpieza de los datos contractuales en Python .....	23
<b>Figura 7</b> Depuración de duplicados, nulos y tratamiento de valores atípicos.....	23
<b>Figura 8</b> Exportación y vista previa de la base de datos depurada .....	24
<b>Figura 9</b> Inicio segunda fase análisis exploratorio de la contratación .....	24
<b>Figura 10</b> Totales adjudicados por modalidad y dependencia.....	24
<b>Figura 11</b> Resultados del análisis exploratorio de la contratación.....	25
<b>Figura 12</b> Evolución mensual del monto adjudicado en procesos contractuales.....	26
<b>Figura 13</b> Monto adjudicado por modalidad de contratación .....	27
<b>Figura 14</b> Distribución del monto adjudicado por dependencia municipal .....	28
<b>Figura 15</b> Detección automática de columnas relevantes para la predicción .....	29
<b>Figura 16</b> Agregación mensual y aplicación del modelo Prophet .....	30
<b>Figura 17</b> Proyección mensual de montos adjudicados generada por el modelo Prophet.....	31
<b>Figura 18</b> Pronóstico comparativo de montos adjudicados (Modelo Prophet).....	32
<b>Figura 19</b> Detección de anomalías Isolation Forest.....	33
<b>Figura 20</b> Puntaje de anomalía vs. monto adjudicado (Isolation Forest).....	34

## Resumen

La Alcaldía municipal de Puerto Rico, Meta, gestiona anualmente un gran volumen de procesos contractuales, fundamentales para el desarrollo institucional y el uso eficiente de los recursos públicos. Sin embargo, la variabilidad en los montos adjudicados y la falta de herramientas analíticas avanzadas dificultan la planeación presupuestal y la detección temprana de irregularidades. Actualmente, el análisis de la información se limita a reportes descriptivos que no permiten anticipar riesgos administrativos o financieros, frente a esta situación se propone el desarrollo de un algoritmo en Python basado en técnicas de Machine Learning, orientado a la predicción de valores contractuales y la detección de anomalías en los procesos de contratación. Se implementarán modelos de series de tiempo como Prophet para proyectar tendencias y Isolation Forest para identificar comportamientos atípicos. Además, se incorporará la metodología SHAP para lograr garantizar la interpretabilidad de los resultados. La propuesta contempla seis fases: consolidación y limpieza de datos, análisis exploratorio de tendencias, modelado predictivo, detección de anomalías, explicabilidad e integración de un dashboard interactivo. Con ello, se busca fortalecer la planeación presupuestal y eficiencia en la gestión contractual de la Alcaldía de Puerto Rico, Meta.

**Palabras Clave:** Algoritmo, Contractuales, Detección, Anomalías, Análisis

### **Abstract**

The municipal government of Puerto Rico, Meta, manages a large volume of contractual processes each year, which are fundamental for institutional development and the efficient use of public resources. However, the variability in the amounts awarded and the lack of advanced analytical tools make budget planning and early detection of irregularities difficult. Currently, information analysis is limited to descriptive reports that do not allow for the anticipation of administrative or financial risks. In view of this situation, we propose the development of a Python algorithm based on machine learning techniques, aimed at predicting contract values and detecting anomalies in the contracting process. Time series models such as Prophet will be implemented to project trends and Isolation Forest to identify atypical behavior. In addition, the SHAP methodology will be incorporated to ensure the interpretability of the results. The proposal consists of six phases: data consolidation and cleaning, exploratory trend analysis, predictive modeling, anomaly detection, explainability, and integration of an interactive dashboard. The aim is to strengthen budget planning and efficiency in the contractual management of the Mayor's Office of Puerto Rico, Meta.

***Key Word:*** Algorithm, Contractual, Detection, Anomalies, Analysis.

## Introducción

La contratación pública se destaca por ser un proceso clave para garantizar el buen desarrollo de proyectos que impactan directamente a la comunidad. En el municipio de Puerto Rico, Meta, la alcaldía municipal, por medio de su área de contratación, realiza anualmente una gestión con un número considerable de procesos contractuales que presenta una alta variabilidad en sus montos económicos y complejidad administrativa.

Sin embargo, actualmente la gestión de procesos contractuales enfrenta dificultades en el manejo de la información, ya que el análisis se realiza principalmente a través de reportes descriptivos y hojas de cálculo, limitando la capacidad de detectar irregularidades o anticiparse a posibles tendencias, esta situación ha generado un gran obstáculo para llevar a cabo una planeación presupuestal adecuada, aumentando las probabilidades de riesgos financieros y administrativos. Ante esta problemática se plantea el desarrollo de un algoritmo en Python, basado principalmente en técnicas de Machine Learning, orientado a la predicción de los montos adjudicados y la detección de anomalías en los procesos contractuales, para ejecutarlo se emplearán modelos de series de tiempo como prophet, permitiendo proyectar valores económicos y el algoritmo Isolation Forest, para la identificación de comportamientos atípicos, así mismo se integrará la metodología SHAP con el fin de garantizar la interpretación de los resultados por parte de los responsables en el área de contratación.

Por lo tanto, el desarrollo de este sistema busca dotar a la alcaldía del municipio de Puerto Rico, Meta con una herramienta predictiva y de alerta temprana que ayude al fortalecimiento de la planeación presupuestal, mejorando la capacidad de gestión, promoviendo la transparencia y eficiencia en el manejo de los procesos contractuales municipales.

**Título del proyecto**

Algoritmo en Python para predecir y detectar anomalías en procesos contractuales de la alcaldía de puerto rico, meta, mediante técnicas avanzadas de machine Learning

### **Planteamiento del caso**

En el municipio de Puerto rico, Meta , la contratación pública es un eje fundamental para llevar a cabo la ejecución de proyectos de inversión y una gestión de recursos públicos adecuada, sin embargo actualmente la gestión de los procesos contractuales presentan limitaciones de planeación y análisis, debido a que la información se administra principalmente en hojas de cálculo, aunque cumplen su función básica de registro, resultan insuficientes al momento de proyectar tendencias o identificar patrones, generando retrasos en la toma de decisiones y dificultando la planeación presupuestal a mediano y largo plazo.

Esta problemática impacta directamente a la eficiencia y transparencia en la administración pública, debido a la alta variabilidad en montos adjudicados, retrasos en la ejecución contractual y inconsistencias en auditorías internas y externas, evidenciando la falta de un sistema más avanzado que integre herramientas predictivas, sin embargo en la legislación colombiana, la contratación estatal está regulada por la Ley 80 de 1993 y la Ley 1150 de 2007, las cuales establecen principios fundamentales como eficiencia, transparencia y responsabilidad en el uso de los recursos públicos, así como la implementación del sistema Electrónico de Contratación Pública(SECOP), plataforma digital que centraliza la información contractual, a pesar de contar con este marco normativo tan robusto, persisten limitaciones en municipios pequeños presentando dificultades por falta de herramientas tecnológicas de apoyo en la gestión.

Se plantea la necesidad de desarrollar un algoritmo en Python con técnicas de Machine Learning, permitiendo mejorar la planeación y control de los procesos contractuales, haciendo uso de un modelo de serie de tiempo como Prophet para facilitar la predicción de montos adjudicados, mientras que Isolation Forest nos permitirá detectar anomalías de forma eficaz, complementando con la metodología SHAP para lograr una mayor interpretación en los resultados, buscando mejorar la eficiencia en la gestión de procesos contractuales.

### **Marco de referencia**

La contratación pública corresponde a un proceso fundamental en la gestión del estado, permitiendo la ejecución de proyectos de inversión y la prestación de servicios a la ciudadanía, sin embargo “representa un poderoso medio de inversión de fondos públicos. Por lo tanto, es fundamental mejorar dos aspectos principales en el ámbito de la transparencia y el seguimiento de todo el ciclo de inversión y contratación”. (Siciliani et al, 2023), la falta de herramientas avanzadas de análisis de datos en la administración pública ha generado retrasos en la toma de decisiones, afectando la eficiencia de los procesos contractuales, es por esto que la transparencia en la contratación pública es un aspecto crítico y la falta de un sistema de información limita la capacidad de control sobre el ciclo de contratación, La implementación de la integridad algorítmica en la contratación pública, en particular a través del análisis predictivo, se reconoce cada vez más como un enfoque eficaz para mejorar la supervisión y mitigar los riesgos. (Ayobami et al, 2023)

El uso de herramientas como modelos de serie de tiempo, algoritmo de detección de anomalías y técnicas para interpretación de resultados permitirán identificar aquellos patrones de comportamientos atípicos, prediciendo posibles irregularidades y optimizando la gestión de los recursos públicos, teniendo un impacto directo en la planeación presupuestal, permitiendo estimaciones más precisas en los montos adjudicados, contribuyendo a una reducción de riesgos financieros y mejorando la eficiencia administrativa de las entidades públicas, por lo tanto fortalecerá la capacidad de planeación al integrar metodologías de interpretación de resultados en la contratación pública aumentando la transparencia, esta iniciativa no solo promete agilizar los procesos de toma de decisiones, sino que también abre el camino hacia la innovación y la eficiencia en un sector tradicionalmente fragmentado en cuanto a datos. (Mishra et al, 2024), Las herramientas de aprendizaje automático parecen especialmente adecuadas para este fin. De hecho, las investigaciones existentes ya han explorado esta posibilidad y han intentado predecir irregularidades en los contratos de desarrollo internacional. (Gallego et al, 2021)

Actualmente hacen uso de Machine Learning para detectar ineficiencias como sobrecostos y índices de riesgo para predecir irregularidades en el proceso de contratación. Estos dos componentes cubren aspectos complementarios del proceso de contratación, teniendo en cuenta tanto el despilfarro activo como el pasivo, y ayudan a los organismos de supervisión a priorizar las investigaciones y asignar los recursos. (Salazar et al, 2024), implementar técnicas de interpretación

de resultados como SHAP facilitará a los responsables del área de contratación comprender las variables que más afectan en la ejecución de contratos, esta técnica proporciona una forma más clara y basada en principios para medir la importancia de las características con el fin de mejorar la transparencia, la equidad y el rendimiento de algoritmos, (Bowen, D., & Ungar, L., 2020), fortaleciendo la capacidad de tomar decisiones informadas en la administración pública.

La integración de Python como una herramienta de desarrollo en procesos de contratación pública permite automatizar la recolección, limpieza y análisis de datos, facilitando la generación de modelos predictivos, Python es considerado uno de los lenguajes de programación más populares para la ciencia de datos. Por ello, cuenta con un elevado número de útiles librerías complementarias y una comunidad de código abierto. (Raschka & Mirjalili, 2020), además tiene una capacidad de compatibilidad con todos los lenguajes de programación, permitiendo transferir y ejecutar en cualquier plataforma, es fácil de leer y entender el código, ofreciendo una gran variedad de bibliotecas, algunas de las cuales utilizan excelentes herramientas de visualización.

Los modelos predictivos permiten evaluar rápidamente múltiples opciones, especialmente cuando se pueden generar representaciones visuales de los resultados, lo que facilita supervisar visualmente las decisiones algorítmicas y tomar decisiones más precisas y fundamentadas en la administración pública. (Quevedo-Piratova et al, 2024), Además los avances tecnológicos en almacenamiento de datos han impulsado el uso de sensores de alta velocidad que permitan la observación continua del comportamiento de sistemas complejos, lo cual ha facilitado el camino para el diseño de enfoques eficientes de aprendizaje automático no supervisado orientado a la detección de anomalías (Xu et al, 2023).

La transformación digital de la contratación pública ha sido un proceso espontáneo, siendo impulsado por la tecnología a un ritmo diferente en todo el mundo, lo que ha exigido adaptarse a leyes administrativas nacionales y hacer uso de instrumentos internacionales destinados a estandarizar la contratación pública a nivel mundial (Górski, 2023), esta evolución ha favorecido en la implementación de herramientas tecnológicas avanzadas en la gestión pública, entre ellas el lenguaje de programación Python que permite automatizar procesos de análisis potenciando el uso de los modelos predictivos aplicados en tomas de decisiones en contratación estatal, de esta manera el análisis de datos en las entidades públicas desarrolla un papel estratégico, ya que no solo se encarga de fortalecer la transparencia en los procesos contractuales, sino que también facilita el cumplimiento de los objetivos de desarrollo sostenible (Pérez Astonitas et al, 2021).

Como parte de esta evolución la contratación pública electrónica (e-procurement) se reconoce como una medida eficaz para fortalecer la transparencia y la documentación contractual, con el fin de transformar procedimientos sin necesidad de alterar de manera radical los marcos regulatorios. Su impacto se refleja en la generación de información de gestión, y la eficiencia en adquisiciones rutinarias y complejas, ofreciendo ahorros estimados de entre el 5 % y el 8 % del valor de la contratación (Hochstetter et al, 2023). Los procesos de administración electrónica se han consolidado como una línea estratégica de inversión para garantizar la transparencia en la contratación y adquisición de bienes y servicios públicos, contribuyendo al fortalecimiento del gobierno electrónico y al desarrollo sostenible de la gestión pública. (Cobo Navarrete, I., 2022).

La Industria 4.0 ha fortalecido esta transformación, al impulsar la implementación de tecnologías emergentes en la contratación pública proporcionando un papel importante en la modernización del sector. (Ramos-Rivadeneira, D. X., & Jiménez-Toledo, J. A., 2024), estas innovaciones además de optimizar los procesos administrativos y operativos, también ayudan a reforzar la eficiencia y trazabilidad en la gestión estatal, su integración en los sistemas de e-procurement impulsa la proyección de una administración pública más competitiva, en este mismo aspecto la incorporación de tecnologías digitales como Big Data y Inteligencia Artificial permiten procesar volúmenes de datos de información, identificar patrones de riesgo mejorando la toma de decisiones en tiempo real, fortaleciendo la gestión administrativa en los procesos de contratación (García Rodríguez, M. J. G. 2022), de este modo la necesidad de detectar las ineficiencias de manera temprana en la contratación pública se justifica, dado a las inconsistencias en diferentes etapas del ciclo contractual indicando niveles bajos de solidez institucional (Rodríguez Arévalo, S., 2021).

En Colombia la contratación estatal está regulada por la Ley 80 de 1993, ley tiene por objeto disponer las reglas y principios que rigen los contratos de las entidades estatales. Estos criterios se han consolidado por la Ley 1150 de 2007 y el Decreto 1082 de 2015, los cuales introdujeron mecanismos orientados a la selección objetiva de contratistas y a la optimización de los procesos. (Genes et al, 2021). Sin embargo, informes de la (Organisation for Economic Co-operation and Development (OECD), 2021) determinan, que a pesar de contar con un marco normativo sólido, Colombia actualmente enfrenta dificultades en la aplicación práctica de estas disposiciones debido a la fragmentación de los sistemas de información y la escasa capacidad analítica de los entes territoriales.

**Preguntas de reflexión**

¿De qué manera la aplicación de herramientas de análisis de datos y modelos de Machine Learning puede fortalecer la transparencia y eficiencia en los procesos contractuales municipales?

¿Qué beneficios aporta el uso de modelos de series de tiempo, como Prophet, en la planeación presupuestal y la proyección de los montos adjudicados en la contratación pública?

¿Por qué es fundamental incorporar técnicas de explicabilidad, como SHAP, para garantizar la comprensión y confianza en los resultados generados por los modelos predictivos aplicados a la gestión contractual?

## Objetivos

### Objetivo General

Desarrollar un algoritmo en Python basado en técnicas de Machine Learning que permita predecir valores contractuales y detectar anomalías en los procesos contractuales de la Alcaldía de Puerto Rico, Meta, optimizando la planeación presupuestal y eficiencia administrativa.

### Objetivos específicos

Implementar un modelo de series de tiempo Prophet para proyectar los montos adjudicados en contratos, generando estimaciones que respalden la planeación presupuestal.

Ejecutar el algoritmo de detección de anomalías Isolation Forest para identificar procesos contractuales con valores atípicos o inconsistencias en montos, fechas y modalidades, fortaleciendo los mecanismos de control preventivo.

Incorporar técnicas de explicabilidad con SHAP (SHapley Additive exPlanations) que permitan a los funcionarios del área de contratación interpretar los resultados de los modelos y comprender las variables más influyentes en las predicciones, promoviendo una toma de decisiones informada

## Marco Metodológico

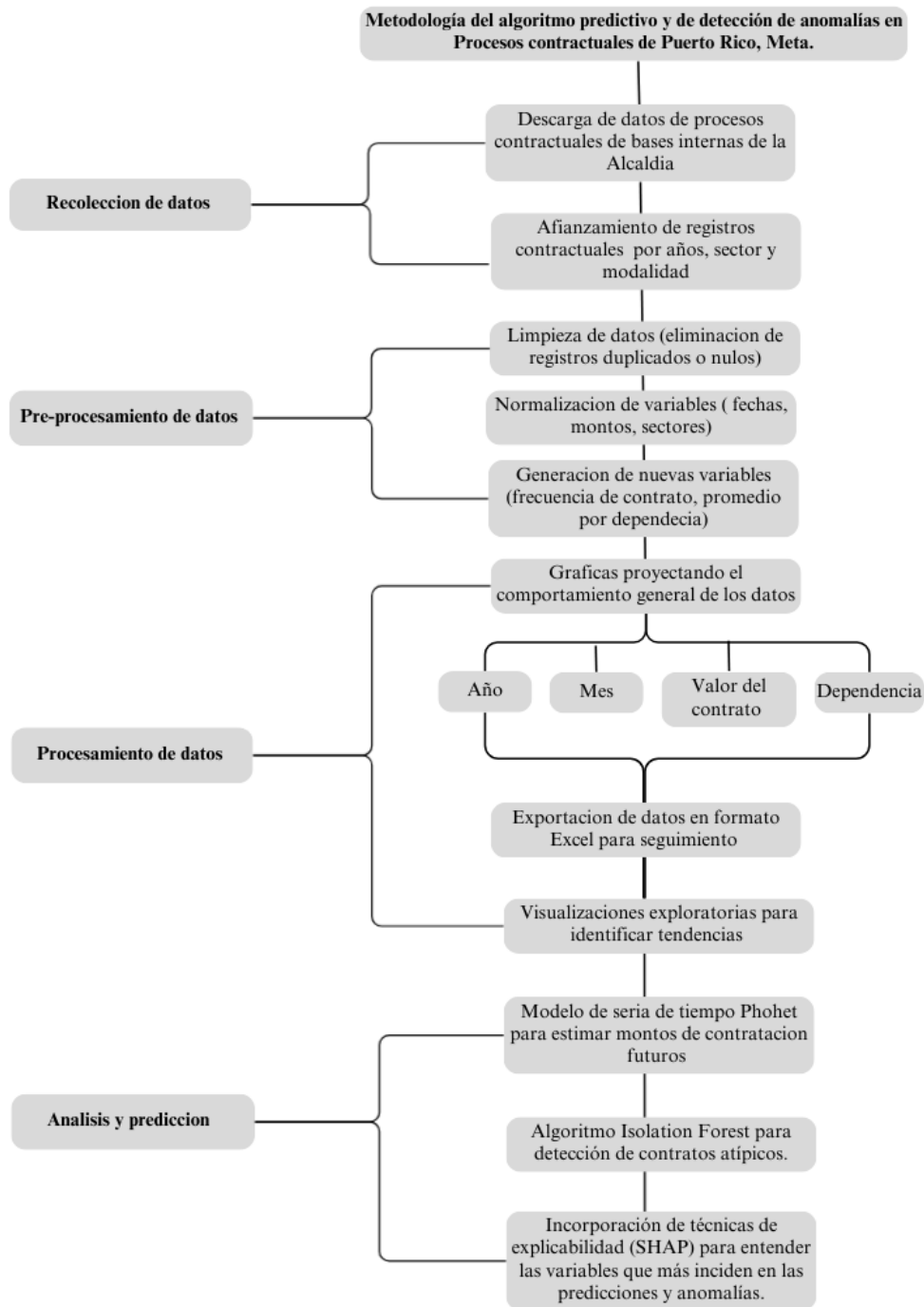
### Tipo de estudio

Este proyecto se llevará a cabo por medio de un estudio con enfoque cuantitativo, buscando desarrollar un algoritmo en Python basándonos en técnicas de Machine Learning favoreciendo significativamente la gestión interna de la Alcaldía de Puerto Rico, Meta. El diseño es de tipo descriptivo y analítico debido a que caracteriza procesos contractuales, a su vez también predictivo porque proyecta tendencias y detecta anomalías para su desarrollo se contemplan cinco fases:

1. **Recolección y limpieza de datos contractuales:** En esta etapa se consolida la información de los procesos contractuales provenientes de las bases internas de la Alcaldía de Puerto Rico, Meta, después, se realiza un proceso de depuración de datos, eliminando registros duplicados, corrigiendo valores atípicos y estandarizando los campos, con el propósito de garantizar su coherencia y fiabilidad para el análisis. Esta fase constituye la base técnica del proyecto, asegurando que las siguientes etapas se fundamenten en datos limpios.
2. **Análisis exploratorio de tendencias:** Una vez organizada la base de datos, se realiza un análisis exploratorio para comprender el comportamiento histórico de la contratación en el municipio. Se examinan variables como la modalidad del contrato, contratación directa, licitación, selección abreviada, el monto adjudicado por dependencia y la evolución mensual de los procesos. Mediante el uso de gráficos y resúmenes estadísticos, se identifican patrones de concentración y estacionalidades, como picos de contratación a finales de año. Estos hallazgos permiten establecer una línea base de comportamiento, esencial para la calibración del modelo predictivo Prophet en la siguiente fase.
3. **Aplicación de modelo de predicción Prophet:** En esta fase se aplica el modelo de series de tiempo Prophet para analizar la evolución de los montos adjudicados y proyectar su comportamiento futuro. Este modelo permite capturar tendencias y estacionalidades características de la contratación pública, como los incrementos en periodos específicos del año. El uso de Prophet permite generar proyecciones confiables, que apoyan la planeación presupuestal y contribuyen a anticipar escenarios de sobrecostos o subejecución de recursos, proporcionando una herramienta de apoyo para la toma de decisiones en la gestión administrativa.

- 4. *Implementación de algoritmo de detección de anomalías Isolation forest:*** Esta fase tiene como finalidad detectar procesos contractuales atípicos frente al patrón general de comportamiento identificado. Utilizando el algoritmo Isolation Forest, se analizan variables como los montos adjudicados, los plazos contractuales y las modalidades de selección para identificar posibles irregularidades o inconsistencias. La detección temprana de anomalías proporciona un mecanismo de control preventivo, que permite alertar sobre riesgos financieros o administrativos, fortaleciendo la transparencia y el control interno en la gestión contractual municipal.
- 5. *Incorporación de técnicas de explicabilidad SHAP:*** En la fase final se integran las técnicas de explicabilidad SHAP (SHapley Additive exPlanations), con el propósito de interpretar los resultados obtenidos de los modelos predictivos y de detección de anomalías. Esta metodología permite identificar la influencia de las variables más relevantes, como el valor del contrato, el tipo de proceso, la dependencia ejecutora o la vigencia, sobre las predicciones generadas por el algoritmo. De este modo, los responsables del área de contratación pueden comprender las razones detrás de cada predicción o señalamiento de anomalía, reforzando la confianza en el modelo y promoviendo la toma de decisiones fundamentadas y transparentes, en concordancia con los principios de eficiencia y responsabilidad de la contratación estatal colombiana.

**Figura 1** Metodología del algoritmo predictivo y de detección de anomalías en procesos contractuales de Puerto Rico, Meta.



*Nota.* Representación del flujo metodológico adaptado para la predicción y detección de anomalías en procesos contractuales.

**Figura 2** Cronograma de actividades metodológicas para el proyecto de predicción y detección de anomalías contractuales.

FASE	ACTIVIDADES	AGOSTO				SEPTIEMBRE				OCTUBRE				NOVIEMBRE			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Planificación	Definición de objetivos y alcance del proyecto	■	■	■													
	Revisión base de datos procesos contractuales	■	■	■													
Carga y limpieza de datos	Importación de datos en Python (Google colab)			■	■												
	Limpieza y preprocesamiento de datos			■	■												
Análisis exploratorio	Visualización de datos			■	■	■	■										
	Selección de modelos			■	■	■	■										
Modelo predictivo	Desarrollo de modelo Prophet			■	■	■	■										
	Implementación Isolation Forest					■	■	■	■								
Explicabilidad	Técnicas SHAP para interpretación					■	■	■	■								
Resultados	Desarrollo del dashboard y consolidación de hallazgos							■	■	■	■						
Documentación	Redacción del documento final							■	■	■	■						
	Revisión con director y ajustes							■	■	■	■						
Entrega	Entrega de documento final							■	■	■	■						
	Sustentación del trabajo de grado														■	■	

Nota. El cronograma muestra la distribución semanal de las fases metodológicas y actividades principales desarrolladas en el marco del proyecto de grado sobre predicción y detección de anomalías en procesos contractuales de la Alcaldía de Puerto Rico, Meta.

### Narración del caso

El presente estudio se desarrolló en el marco de la gestión contractual de la Alcaldía del municipio de Puerto Rico, Meta, institución responsable de la planeación, adjudicación y seguimiento de los procesos de contratación pública, inicié identificando aquellas limitaciones que existen actualmente en la planeación presupuestal, se encontraron asociadas a la ausencia de herramientas predictivas, dependiendo de reportes descriptivos básicos, lo que generaba retrasos en la toma de decisiones y detección de posibles anomalías internas.

En este caso se realizó un proceso de diagnóstico en la unidad de contratación, dirigida por Cristian Alejandro Perez Muñoz, Jefe de la dependencia, iniciando por una recolección de información sobre métodos utilizados actualmente para consolidar datos presupuestales y de adjudicación, así como tiempos en tomas de decisión, evidenciando la necesidad de fortalecer la capacidad para anticipar los riesgos administrativos

Posteriormente se analizó las limitaciones existentes en la planeación presupuestal, formulando una propuesta de mejora orientada a automatizar los procesos internos de esta dependencia, con el fin de implementar una herramienta de análisis predictivo, incorporando modelos de proyección de tendencias presupuestales y mecanismos de detección de anomalías en los procesos contractuales, la propuesta fue estructurada de manera integral, considerando tanto el componente técnico como el organizacional, incluyendo la capacitación del personal de la Unidad

de Contratación en el uso de modelos de predicción y visualización de datos. Se buscó que la herramienta no solo funcionara como un software, sino como un mecanismo de gestión del conocimiento, capaz de articular la planeación, el seguimiento y la evaluación de los procesos en tiempo real.

Esta iniciativa fue presentada al área de contratación y aceptada para su implementación dentro de la gestión contractual de la Alcaldía, estableciendo un esta iniciativa fue presentada al área de contratación y aceptada para su implementación permitiendo la medición de los modelos propuestos evaluando su impacto en la toma de decisiones en la administración municipal.

Primero se realizó la importación de las librerías esenciales para el algoritmo: numpy para cálculos numéricos, pandas para el manejo de datos, y matplotlib.pyplot para la generación de gráficos. También se incluyeron warnings para la gestión de rutas y el control de advertencias. Se configuró el parámetro figure.autolayout para optimizar la visualización de gráficas y se definió una originaridad aleatoria (RANDOM\_STATE = 42) que asegura la reproducibilidad de los resultados.

**Figura 3** Configuración inicial del entorno de trabajo en Python

**ALGORITMO GESTIÓN CONTRACTUAL ALCALDÍA PUERTO RICO, META**

```
[31]
✓ 0 s
import warnings, os, numpy as np, pandas as pd, matplotlib.pyplot as plt
warnings.filterwarnings("ignore")
plt.rcParams.update({'figure.autolayout': True})
RANDOM_STATE = 42
```

*Nota.* Se muestran las librerías importadas Fuente: Elaboración propia con base en ejecución en *Google Colaboratory* (2025).

En esta etapa se realice la carga del archivo PROCESOS\_CONTRACTUALES.xlsx, que contiene la información de los procesos contractuales de la Alcaldía de Puerto Rico, Meta. El archivo se importó al entorno de Google Colaboratory, quedando almacenado en la carpeta de trabajo (/content/) para su posterior procesamiento

**Figura 4** Carga del archivo de procesos contractuales en Google Colaboratory

```
[32]
✓ 13 s
from google.colab import files
print("/content/PROCESOS_CONTRACTUALES.xlsx")
uploaded = files.upload()

/content/PROCESOS_CONTRACTUALES.xlsx
PROCESO...TUALES.xlsx
PROCESOS_CONTRACTUALES.xlsx(application/vnd.openxmlformats-officedocument.spreadsheetml.sheet) - 680168 bytes, last modified: 17/9/2025 - 100% done
Saving PROCESOS_CONTRACTUALES.xlsx to PROCESOS_CONTRACTUALES (1).xlsx
```

*Nota.* Se muestra el procedimiento de importación del archivo PROCESOS\_CONTRACTUALES.xlsx al entorno de ejecución, mediante la instrucción files.upload().

En la siguiente etapa se construyó un diccionario de mapeo de columnas, cuyo objetivo fue estandarizar los nombres originales de la base de datos a variables más claras y consistentes para el análisis. por ejemplo campos como fecha se transformaron en fecha\_adjudicacion, y (D) Valor Financiado como Recursos Propios en monto\_adjudicado, incluyendo además variantes de escritura para las variables de modalidad y tipo de contrato, garantizando así la correcta lectura de los datos independientemente de cómo se encontraban plasmados en el archivo original.

**Figura 5** Definición del mapeo de columnas para la base de datos

```
3] MAPEO_COLUMNAS = {  
0s   'FECHA': 'fecha_adjudicacion',  
      '(D) Valor Financiado Con Recursos Propios': 'monto_adjudicado',  
      # Dos posibles nombres para modalidad y tipo de contrato (usa el que exista)  
      'MODALIDAD': 'modalidad',  
      '(C) Modalidad': 'modalidad',  
      'TIPO DE CONTRATO': 'tipo_contrato',  
      '(C) Tipo De Contrato': 'tipo_contrato',  
      'SECRETARIA': 'dependencia',  
  }
```

*Nota.* Se muestra la definición del mapeo de columnas para unificar los nombres de la base de datos y trabajar con variables estandarizadas.

Iniciamos con la primer fase de las cinco implementadas en el marco metodológico, en esta fase se realice la recolección y limpieza de los datos contractuales, leyendo la base original en Excel y aplicando transformaciones para estandarizar los nombres de las columnas, convertir formatos de fechas y valores numéricos, generando la variable de vigencia a partir del año de adjudicación.

**Figura 6** *Recolección y limpieza de los datos contractuales en Python*

```

[34] ✓ 0 s
# Fase 1: Recolección y limpieza
def cap_outliers_iqr(s: pd.Series) -> pd.Series:
    q1, q3 = s.quantile([0.25, 0.75])
    iqr = q3 - q1
    low, high = q1 - 1.5 * iqr, q3 + 1.5 * iqr
    return s.clip(lower=low, upper=high)

[35] ✓ 1 s
df_raw = pd.read_excel('/content/PROCESOS_CONTRACTUALES.xlsx', header=1)
df = df_raw.copy()

[36] ✓ 0 s
rename_dict = {k: v for k, v in MAPEO_COLUMNAS.items() if k in df.columns}
df.rename(columns=rename_dict, inplace=True)

[37] ✓ 0 s
if 'fecha_adjudicacion' in df.columns:
    df['fecha_adjudicacion'] = pd.to_datetime(df['fecha_adjudicacion'], errors='coerce')

[38] ✓ 0 s
if 'monto_adjudicado' in df.columns:
    df['monto_adjudicado'] = pd.to_numeric(df['monto_adjudicado'], errors='coerce')

[39] ✓ 0 s
if 'plazo_dias' in df.columns:
    df['plazo_dias'] = pd.to_numeric(df['plazo_dias'], errors='coerce')

[40] ✓ 0 s
if 'fecha_adjudicacion' in df.columns:
    df['vigencia'] = df['fecha_adjudicacion'].dt.year

```

*Nota.* Se muestra el proceso de lectura del archivo y la estandarización de nombres de columnas, con la conversión de variables de fecha y numéricas, así como la creación de la variable vigente a partir del año de adjudicación.

Se continua con la depuración de la base de datos eliminando duplicados y gestionando valores nulos mediante imputación con la mediana en variables numéricas y la asignación de “DESCONOCIDO” en categóricas, aplicando el método de rangos intercuartílicos (IQR) para el tratamiento de valores atípicos en los campos de monto adjudicado y plazo en días.

**Figura 7** *Depuración de duplicados, nulos y tratamiento de valores atípicos*

```

41] ✓ 0 s
# Depuración: duplicados y nulos
df = df.drop_duplicates()

42] ✓ 0 s
# Imputaciones simples
if 'monto_adjudicado' in df.columns:
    df['monto_adjudicado'] = df['monto_adjudicado'].fillna(df['monto_adjudicado'].median())

for c in ['modalidad', 'tipo_contrato', 'dependencia', 'vigencia']:
    if c in df.columns:
        df[c] = df[c].fillna('DESCONOCIDO')

43] ✓ 0 s
# Tratamiento de atípicos (IQR) en numéricas clave
for c in ['monto_adjudicado', 'plazo_dias']:
    if c in df.columns:
        df[c] = cap_outliers_iqr(df[c])

```

*Nota.* Se eliminaron registros repetidos, se imputaron valores faltantes y se ajustaron montos extremos con el método IQR.

Completamos la limpieza de la base de datos y se exporó en formato Excel para su almacenamiento. Posteriormente, se visualizó una muestra de las primeras filas con el fin de comprobar la correcta transformación y estandarización de los registros contractuales.

**Figura 8** Exportación y vista previa de la base de datos depurada

```
[44] ✓ 0 s
# Guardar versión limpia en Excel
if 'OUTPUT_DIR' not in locals() and 'OUTPUT_DIR' not in globals():
    OUTPUT_DIR = 'output'
    os.makedirs(OUTPUT_DIR, exist_ok=True)

clean_path = f'{OUTPUT_DIR}/content/PROCESOS_CONTRACTUALES.xlsx'
os.makedirs(os.path.dirname(clean_path), exist_ok=True)
df.to_excel(clean_path, index=False)
print("Fase 1 (limpieza) completada. Base depurada:", clean_path)
display(df.head(3))

Fase 1 (limpieza) completada. Base depurada: output/content/PROCESOS_CONTRACTUALES.xlsx
```

NO	fecha_adjudicacion	Invitación	modalidad	tipo_contrato	Nº CONTRATO	OBJETO	C.C /NIT	CONTRATISTA	dependencia	(D) Valor Financiado Con Otras Fuentes	(D) Valor Adiciones Con Recursos Pregón	monto_adjudicado	(D) Valor Adiciones Otros Recursos	modalidad	tipo_contrato	(C) Nit 0 Cedula Del Contratista	(C) Nombre Del Contratista	(F) Fecha de Liquidación	vigencia		
0	1.0	2025-01-07	CD 001-2025	DIRECTA	PRESTACION DE SERVICIO	180.08.10.001-2025	PRESTACION DE SERVICIOS PROFESIONALES DE ABOGA...	17345310	JHON JAIRO REY ORTIZ	SECRETARIA PRIVADA	...	NaN	NaN	20390000.0	NaN	CONTRATACION DIRECTA	PRESTACION DE SERVICIOS	17345310	JHON JAIRO REY ORTIZ	2025-12-30 00:00:00	2025.0
1	2.0	2025-01-08	CD 002-2025	DIRECTA	PRESTACION DE SERVICIO	180.08.10.002-2025	PRESTACION DE SERVICIOS PROFESIONALES DE ABOGA...	19256608	JOSE ALBERTO LEGUIZAMO VELASQUEZ	SECRETARIA PRIVADA	...	NaN	NaN	20390000.0	NaN	CONTRATACION DIRECTA	PRESTACION DE SERVICIOS	19256608	JOSE ALBERTO LEGUIZAMO VELASQUEZ	2025-07-07 00:00:00	2025.0
2	3.0	2025-01-08	CD 003-2025	DIRECTA	PRESTACION DE SERVICIO	180.08.10.003-2025	PRESTACION DE SERVICIOS DE APOYO A LA GESTIÓN...	1006737629	INGRID VANESA PARRA LUGO	UNIDAD DE CONTRATACION	...	NaN	NaN	20390000.0	NaN	CONTRATACION DIRECTA	PRESTACION DE SERVICIOS	1006737629	INGRID VANESA PARRA LUGO	2025-07-07 00:00:00	2025.0

*Nota.* Se muestra la exportación de la base limpia en formato Excel y una vista de las primeras filas de la tabla resultante.

Posteriormente iniciamos la etapa de la fase dos, generando la serie mensual de montos adjudicados y calculando los totales por modalidad y dependencia, con el fin de organizar la información para su representación gráfica posterior.

**Figura 9** Inicio segunda fase análisis exploratorio de la contratación

```
[45] ✓ 0 s
# Fase 2: Análisis exploratorio (sin interpretación)

# Serie mensual
if 'fecha_adjudicacion' in df.columns and 'monto_adjudicado' in df.columns:
    df['mes'] = df['fecha_adjudicacion'].dt.to_period('M').dt.to_timestamp()
    ts_mensual = df.groupby('mes', as_index=False)['monto_adjudicado'].sum()
else:
    ts_mensual = pd.DataFrame()

# Totales por modalidad y por dependencia
# Drop duplicate 'modalidad' column, keeping the first one
if 'modalidad' in df.columns:
    df = df.loc[:,~df.columns.duplicated()]

top_modalidad = (df.groupby('modalidad', as_index=False)['monto_adjudicado'].sum()
                 .sort_values('monto_adjudicado', ascending=False)) if 'modalidad' in df.columns else pd.DataFrame()

top_dependencia = (df.groupby('dependencia', as_index=False)['monto_adjudicado'].sum()
                   .sort_values('monto_adjudicado', ascending=False)) if 'dependencia' in df.columns else pd.DataFrame()
```

*Nota.* Se construye la serie temporal mensual de los montos adjudicados y se calculan los totales agrupados por modalidad y por dependencia.

Se calcularon los totales adjudicados agrupados por modalidad de contratación y por dependencia, con el fin de identificar la distribución de los recursos en cada categoría.

**Figura 10** Totales adjudicados por modalidad y dependencia

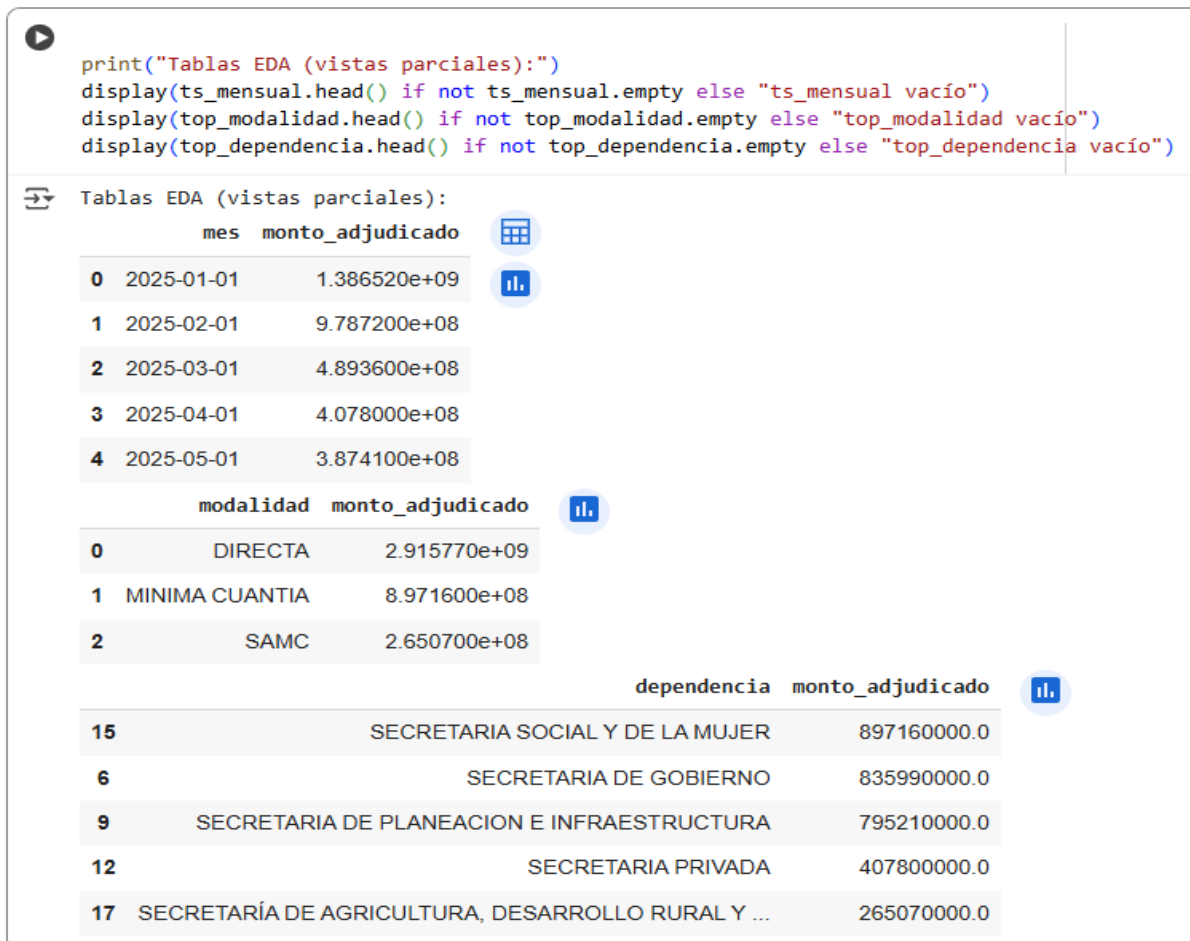
```
[46] ✓ 0 s
# Totales por modalidad y por dependencia
top_modalidad = (df.groupby('modalidad', as_index=False)['monto_adjudicado'].sum()
                 .sort_values('monto_adjudicado', ascending=False)) if 'modalidad' in df.columns else pd.DataFrame()

top_dependencia = (df.groupby('dependencia', as_index=False)['monto_adjudicado'].sum()
                   .sort_values('monto_adjudicado', ascending=False)) if 'dependencia' in df.columns else pd.DataFrame()
```

*Nota.* Se agrupan los registros de contratación para calcular los montos totales adjudicados según la modalidad empleada y la dependencia solicitante.

En esta etapa se generaron tablas parciales con el fin de organizar y resumir la información contenida en la base de datos de contratación. Elaborando una serie temporal mensual que muestra el monto total adjudicado en cada mes, lo cual facilita la identificación de patrones en la evolución del gasto. Posteriormente, se calcularon los totales por modalidad de contratación, lo que permitió reconocer qué mecanismos (como contratación directa, mínima cuantía o selección abreviada) concentran mayores recursos. Finalmente, se construyeron tablas con los totales por dependencia solicitante, que evidencian qué áreas administrativas lideran en volumen de contratación. Estos resultados intermedios representan la base cuantitativa que será utilizada en los gráficos posteriores para facilitar la interpretación visual de las tendencias y distribuciones.

**Figura 11** Resultados del análisis exploratorio de la contratación.



*Nota.* Se muestran las tablas con la serie mensual de montos adjudicados, los totales por modalidad y la distribución por dependencia.

En esta parte se representó gráficamente el monto adjudicado por cada modalidad de contratación. El gráfico de barras permite observar la distribución de los recursos según el mecanismo utilizado, facilitando la comparación entre las diferentes modalidades.

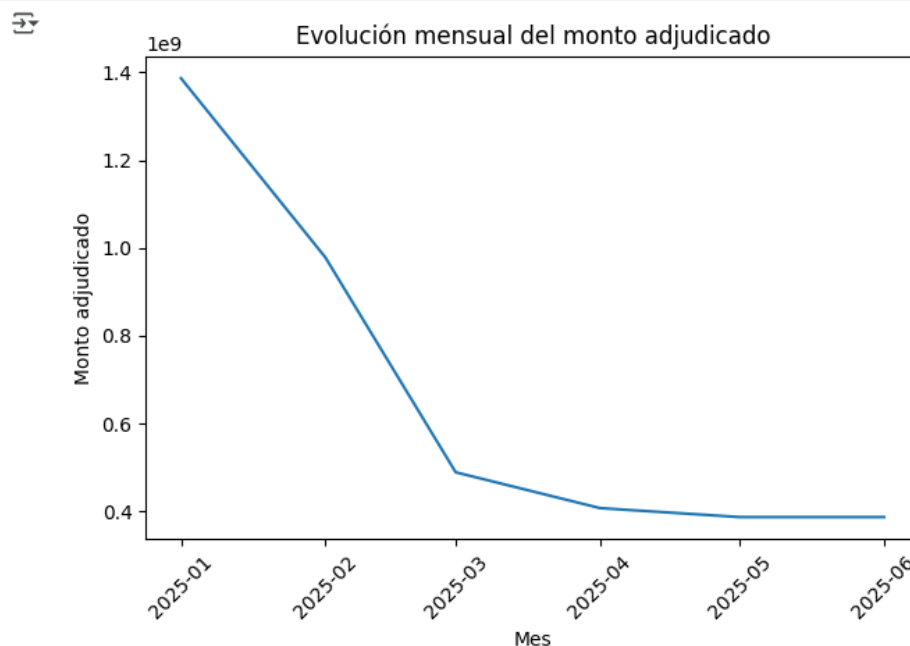
Esto muestra cómo varía el monto adjudicado mes a mes en el año 2025, por ejemplo:

- Enero 2025: 1.386.520.000 COP
- Febrero 2025: 978.720.000 COP
- Marzo 2025: 489.360.000 COP
- Abril 2025: 407.800.000 COP
- Mayo 2025: 387.410.000 COP

Lo significa que existe una tendencia descendente en los montos adjudicados de enero a mayo, con un pico inicial muy alto en enero que luego va bajando cada mes.

**Figura 12** *Evolución mensual del monto adjudicado en procesos contractuales*

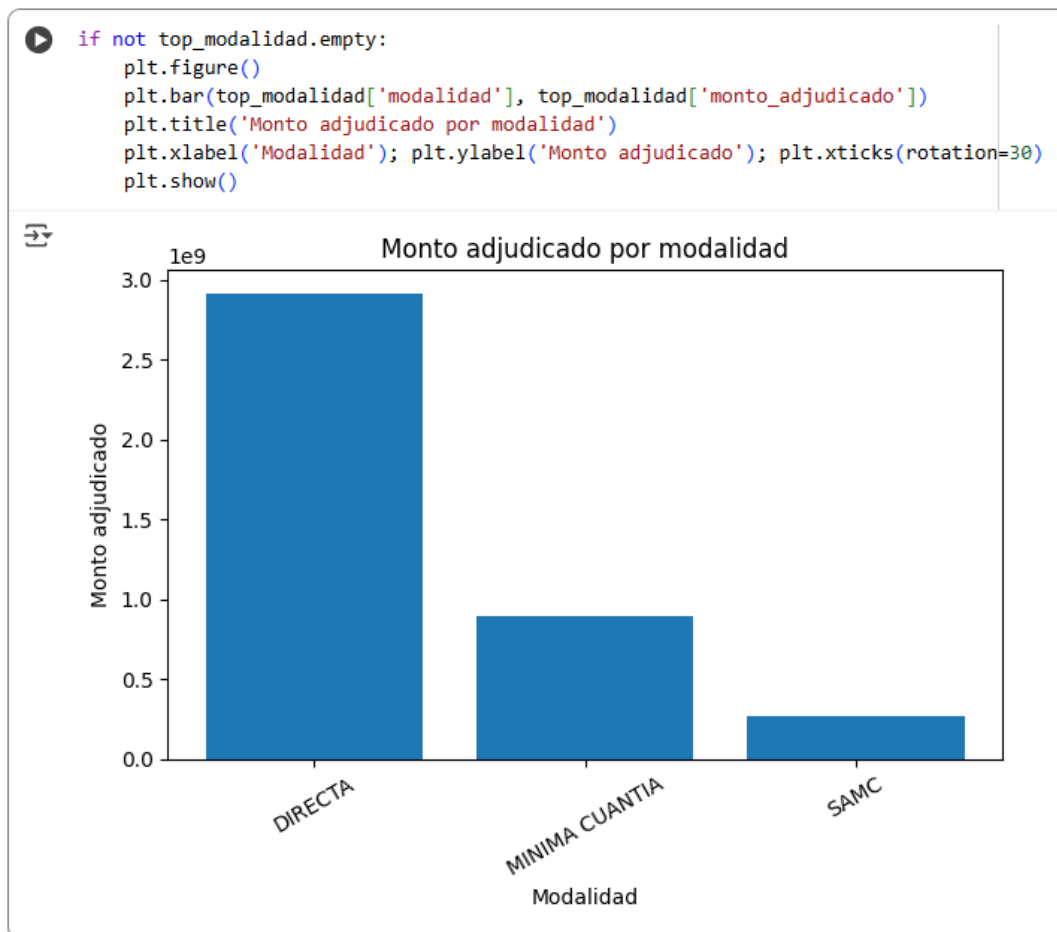
```
if not ts_mensual.empty():
    plt.figure()
    plt.plot(ts_mensual['mes'], ts_mensual['monto_adjudicado'])
    plt.title('Evolución mensual del monto adjudicado')
    plt.xlabel('Mes'); plt.ylabel('Monto adjudicado'); plt.xticks(rotation=45)
    plt.show()
```



*Nota.* El gráfico de líneas muestra la variación mensual de los montos adjudicados en el año 2025. Se observa un pico en enero, seguido de una tendencia descendente hasta estabilizarse en torno a los 380 millones de pesos hacia junio.

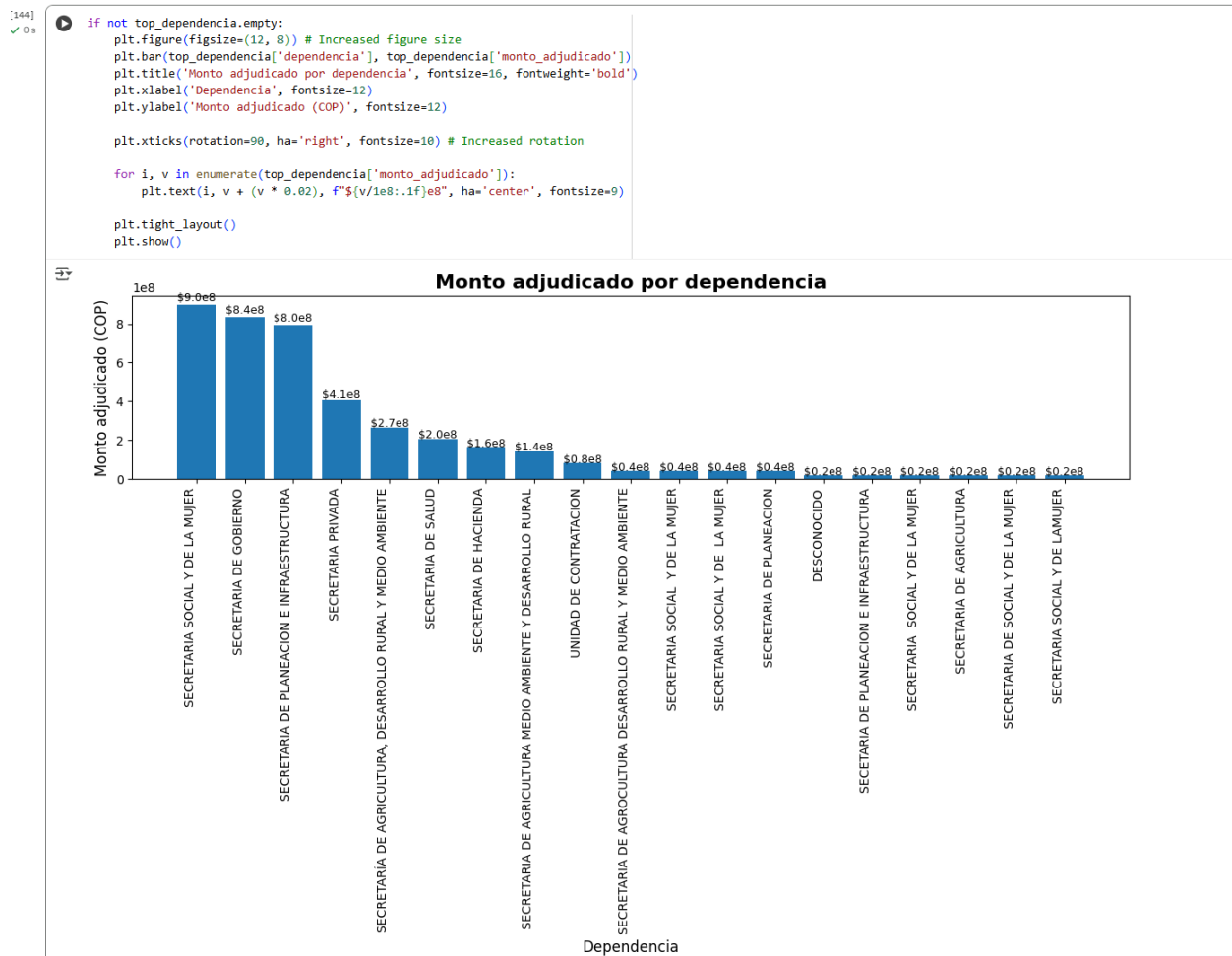
Se observa que hay un descenso pronunciado de enero a marzo, pasando de 1.400 millones a 489 millones, se redujo en casi 65%, por otro lado desde el mes marzo en adelante, los montos adjudicados siguen cayendo pero de manera más suave, hasta estabilizarse alrededor de los 380 millones. Esto podría reflejar que en el inicio del año se adjudican más contratos grandes (enero), y luego la contratación entra en un ritmo más bajo y estable.

**Figura 13** *Monto adjudicado por modalidad de contratación*



*Nota.* Gráfico de barras que muestra la suma de los montos adjudicados agrupados según la modalidad de contratación.

Mediante el gráfico de barras se logra determinar que la modalidad directa se concentra con el mayor monto adjudicado, cercano a 3.0e9 (tres mil millones de pesos), La modalidad Mínima Cuantía ocupa el segundo lugar, pero con una diferencia de aproximadamente 2.1 mil millones de pesos.

**Figura 14** Distribución del monto adjudicado por dependencia municipal

Nota. Gráfico de barras que muestra la suma de los montos adjudicados agrupados según la modalidad de contratación.

El gráfico muestra el monto total adjudicado en pesos colombianos (COP) por cada dependencia del municipio de Puerto Rico, Meta, ordenadas de mayor a menor según el valor total de adjudicación.

- Las primeras barras corresponden a las dependencias con mayor contribución presupuestal, tales como la Secretaría de la Mujer, Gobierno e Infraestructura, evidenciando una concentración significativa del gasto público en estas áreas.
- Las dependencias con menores montos adjudicados aparecen al final del eje X, reflejando baja actividad contractual o proyectos de menor magnitud económica.

Este paso nos permitió identificar tendencias de asignación de recursos, prioridades de inversión municipal y posibles áreas donde podría optimizarse la distribución del presupuesto.

Se continuó con la aplicación del modelo de predicción Prophet, con el propósito de analizar la evolución de los montos adjudicados y proyectar su comportamiento futuro.

**Figura 15** Detección automática de columnas relevantes para la predicción

```
1] #Fase 3: Predicción (Prophet con fallback SARIMAX)
2]
def detectar_columnas(df):
    cols = list(df.columns)
    low = {c.lower(): c for c in cols}

    # candidatos fecha
    fecha_keys = [
        'fecha', 'mes', 'fechacontrato', 'fecha_contrato', 'fecha_adjudicacion',
        'fecha_firma', 'fechafirma', 'fecha_inicio', 'fecha_publicacion'
    ]
    fecha_cands = [low[k] for k in fecha_keys if k in low]
    # incluir columnas datetime
    for c in cols:
        if np.issubdtype(df[c].dtype, np.datetime64) and c not in fecha_cands:
            fecha_cands.append(c)
    fecha_col = fecha_cands[0] if fecha_cands else cols[0]

    # candidatos monto
    monto_keys = [
        'monto', 'valor', 'monto_adjudicado', 'valor_adjudicado', 'valor_total',
        'valortotal', 'valorcontrato', 'valor_contrato', 'cuantia', 'cuantía', 'presupuesto'
    ]
    monto_cands = [low[k] for k in monto_keys if k in low]
    # columnas numéricas como respaldo
    num_cols = [c for c in cols if pd.api.types.is_numeric_dtype(df[c])]
    monto_col = monto_cands[0] if len(monto_cands) else (num_cols[0] if num_cols else cols[1])

    return fecha_col, monto_col
```

Esta parte del código se encarga de analizar los nombres de las variables del DataFrame y los compara con una lista de términos comunes en los registros de contratación pública, tales como `fecha_adjudicación`, `mes`, `monto_adjudicado` o `valor_total`. En caso de no encontrar coincidencias exactas, el algoritmo selecciona las columnas que poseen formato de fecha o valores numéricos, asegurando la detección correcta incluso en bases con diferentes nomenclaturas.

Como resultado, la función devuelve dos variables: `fecha_col`, que contiene la información temporal, y `monto_col`, que almacena los valores económicos de los contratos. Este paso constituye la base para las etapas posteriores de limpieza, agregación mensual y modelado predictivo, permitiendo estandarizar la información y garantizar la compatibilidad del análisis

**Figura 16** *Agregación mensual y aplicación del modelo Prophet*

```

fecha_col, monto_col = detectar_columnas(df)

tmp = df[[fecha_col, monto_col]].copy()
tmp.columns = ['fecha', 'monto']
tmp['fecha'] = pd.to_datetime(tmp['fecha'], errors='coerce', dayfirst=True)
tmp['monto'] = pd.to_numeric(tmp['monto'], errors='coerce')
tmp = tmp.dropna(subset=['fecha', 'monto'])
tmp = tmp[tmp['monto'] >= 0]

if tmp.empty:
    raise ValueError("No hay filas válidas después de limpiar. Revisa las columnas de fecha y monto.")

ts_mensual = (tmp.groupby(pd.Grouper(key='fecha', freq='MS'))['monto']
               .sum()
               .reset_index()
               .rename(columns={'fecha': 'ds', 'monto': 'y'})
               .sort_values('ds')
               .reset_index(drop=True))

if not ts_mensual.empty:
    full_idx = pd.date_range(ts_mensual['ds'].min(), ts_mensual['ds'].max(), freq='MS')
    ts_mensual = ts_mensual.set_index('ds').reindex(full_idx).fillna(0.0).rename_axis('ds').reset_index()

if len(ts_mensual) < 6:
    raise ValueError(f"La serie mensual es muy corta ({len(ts_mensual)} puntos). Se recomiendan al menos 6-12 meses.")

# Nota: Prophet captura tendencia y estacionalidad anual, muy útil en ciclos presupuestales
from prophet import Prophet
m = Prophet(
    yearly_seasonality=True,
    weekly_seasonality=False,
    daily_seasonality=False,
    interval_width=0.8 )

m.fit(ts_mensual.rename(columns={'ds': 'ds', 'y': 'y'}))

#Horizonte de pronóstico
horizon_months = 12
futuro = m.make_future_dataframe(periods=horizon_months, freq='MS', include_history=True)
fc = m.predict(futuro)

forecast_df = fc[['ds', 'yhat', 'yhat_lower', 'yhat_upper']].copy()
forecast_df.tail(horizon_months)

```

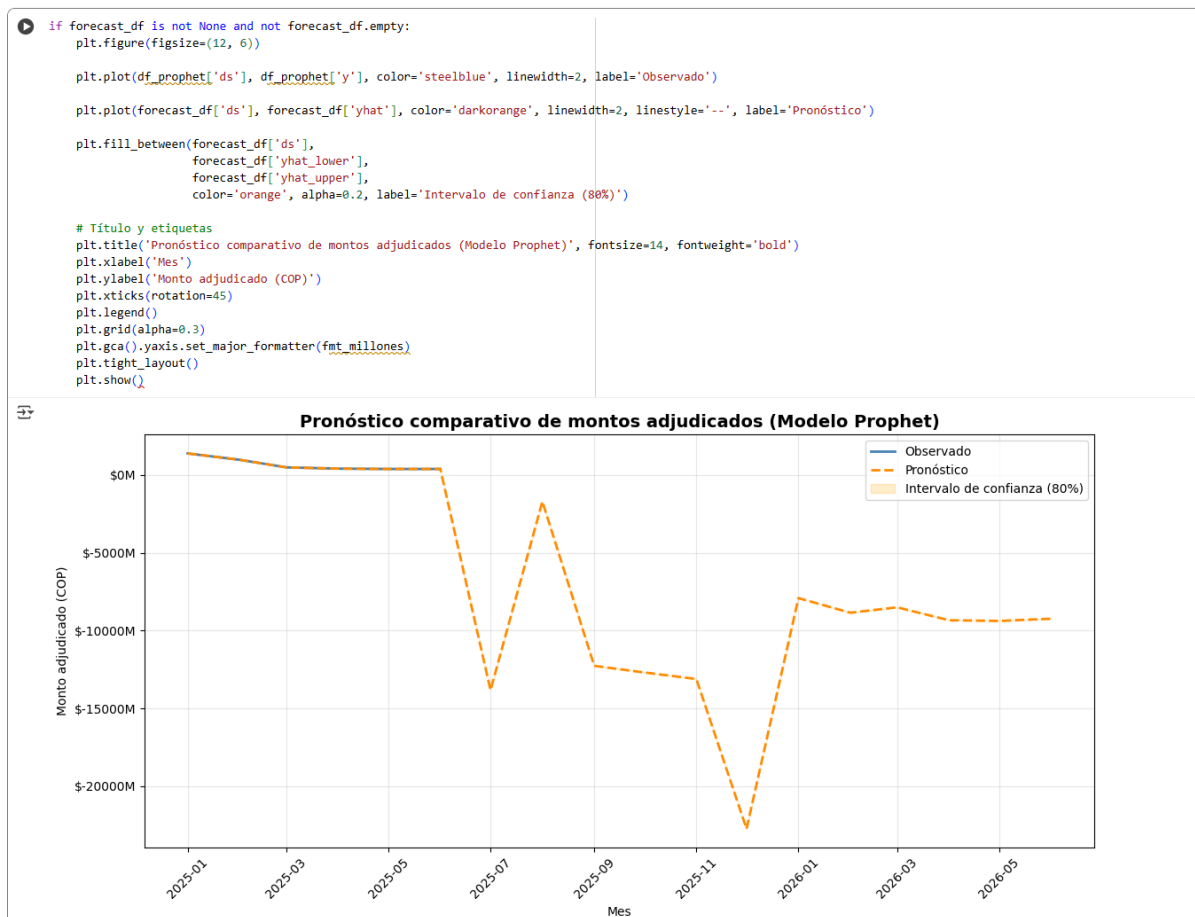
Se presenta el proceso de limpieza, estructuración y modelado de la serie temporal correspondiente a los montos adjudicados. En esta etapa, el código se encarga de transformar los datos que fueron detectados previamente en un formato compatible con el modelo de predicción, se empezó con la depuración de registros nulos o inconsistentes, convirtiendo las fechas y montos a tipos de datos válidos y eliminando valores negativos. Posteriormente, los datos se agrupan por mes, sumando los valores adjudicados para lograr construir una serie temporal mensual. Con el fin de mantener la continuidad del análisis, se rellenan los meses sin adjudicación con valores iguales a cero, lo cual es común en los procesos de contratación pública. Finalmente, se implementa el modelo Prophet, configurado con estacionalidad anual, con el propósito de capturar patrones recurrentes en los ciclos presupuestales. El modelo genera un pronóstico de 12 meses acompañado de un intervalo de confianza del 80%, ofreciendo una proyección confiable del comportamiento futuro de los montos adjudicados.

**Figura 17** Proyección mensual de montos adjudicados generada por el modelo Prophet

	ds	yhat	yhat_lower	yhat_upper
6	2025-07-01	-1.382516e+10	-1.382516e+10	-1.382516e+10
7	2025-08-01	-1.704686e+09	-1.704686e+09	-1.704686e+09
8	2025-09-01	-1.226490e+10	-1.226490e+10	-1.226490e+10
9	2025-10-01	-1.269130e+10	-1.269130e+10	-1.269130e+10
10	2025-11-01	-1.310947e+10	-1.310947e+10	-1.310947e+10
11	2025-12-01	-2.273791e+10	-2.273791e+10	-2.273791e+10
12	2026-01-01	-7.909338e+09	-7.909338e+09	-7.909338e+09
13	2026-02-01	-8.850930e+09	-8.850930e+09	-8.850930e+09
14	2026-03-01	-8.503317e+09	-8.503317e+09	-8.503317e+09
15	2026-04-01	-9.333552e+09	-9.333552e+09	-9.333552e+09
16	2026-05-01	-9.376977e+09	-9.376977e+09	-9.376977e+09
17	2026-06-01	-9.236012e+09	-9.236012e+09	-9.236011e+09

Analizando los valores obtenidos, se evidencio que el modelo proyecta montos negativos a partir de julio de 2025, iniciando con aproximadamente  $-13.800$  millones de pesos) y manteniendo esta tendencia descendente hasta junio de 2026, donde se alcanzan valores cercanos a  $-9.200$  millones de pesos. indicando que según el comportamiento histórico de la serie, el modelo identifica una caída progresiva en el nivel de adjudicación contractual durante los próximos doce meses.

Por lo tanto las proyecciones que presentan valores negativos no implica adjudicaciones reales con signo negativo, sino que refleja una tendencia estadística decreciente, es decir, una pérdida o reducción significativa en la magnitud de los montos adjudicados. Es decir, el modelo está señalando un periodo de contracción en la ejecución presupuestal, donde la dinámica de contratación tiende a disminuir de manera sostenida. es por esto que el modelo Prophet, al capturar patrones de estacionalidad anual, detecta que los incrementos en la contratación se concentran en periodos específicos del año (por ejemplo, los meses previos al cierre fiscal), seguidos por lapsos de menor adjudicación.

**Figura 18** *Pronóstico comparativo de montos adjudicados (Modelo Prophet)*

El gráfico permite evidenciar una tendencia decreciente en la evolución futura de los montos adjudicados, con caídas pronunciadas en determinados meses y una recuperación parcial hacia el primer trimestre de 2026. Este comportamiento sugiere un posible escenario de reducción en la ejecución contractual, lo cual podría asociarse a variaciones estacionales en el ciclo presupuestal o a una disminución de los recursos disponibles, esta proyección obtenida permite anticipar variabilidades en el gasto público y refuerza la utilidad del modelo Prophet como herramienta de apoyo para la planeación y gestión municipal.

Continuando con el desarrollo de la fases propuestas en el marco metodológico, se implementará la cuarta fase en donde se desarrolló el algoritmo de detección de anomalías Isolation Forest, orientado a identificar comportamientos atípicos dentro de los procesos contractuales analizados.

**Figura 19** *Detección de anomalías Isolation Forest*

```
# Detección de anomalías

from sklearn.ensemble import IsolationForest
from sklearn.preprocessing import OneHotEncoder
from sklearn.compose import ColumnTransformer
from sklearn.pipeline import Pipeline
import matplotlib.pyplot as plt

# Definir features (numéricas + categóricas)
features_num = [c for c in ['log_monto', 'plazo_dias', 'anio', 'mes_num'] if c in df_anom.columns]
features_cat = [c for c in ['modalidad', 'tipo_contrato', 'dependencia'] if c in df_anom.columns]
X = df_anom[features_num + features_cat].copy()

# Preprocesamiento
pre = ColumnTransformer(
    transformers=[('onehot', OneHotEncoder(handle_unknown='ignore'), features_cat)],
    remainder='passthrough'
)

# Isolation Forest con contaminación = 5% (marca ~5% como anómalos)
iso = IsolationForest(n_estimators=300, contamination=0.05, random_state=42)

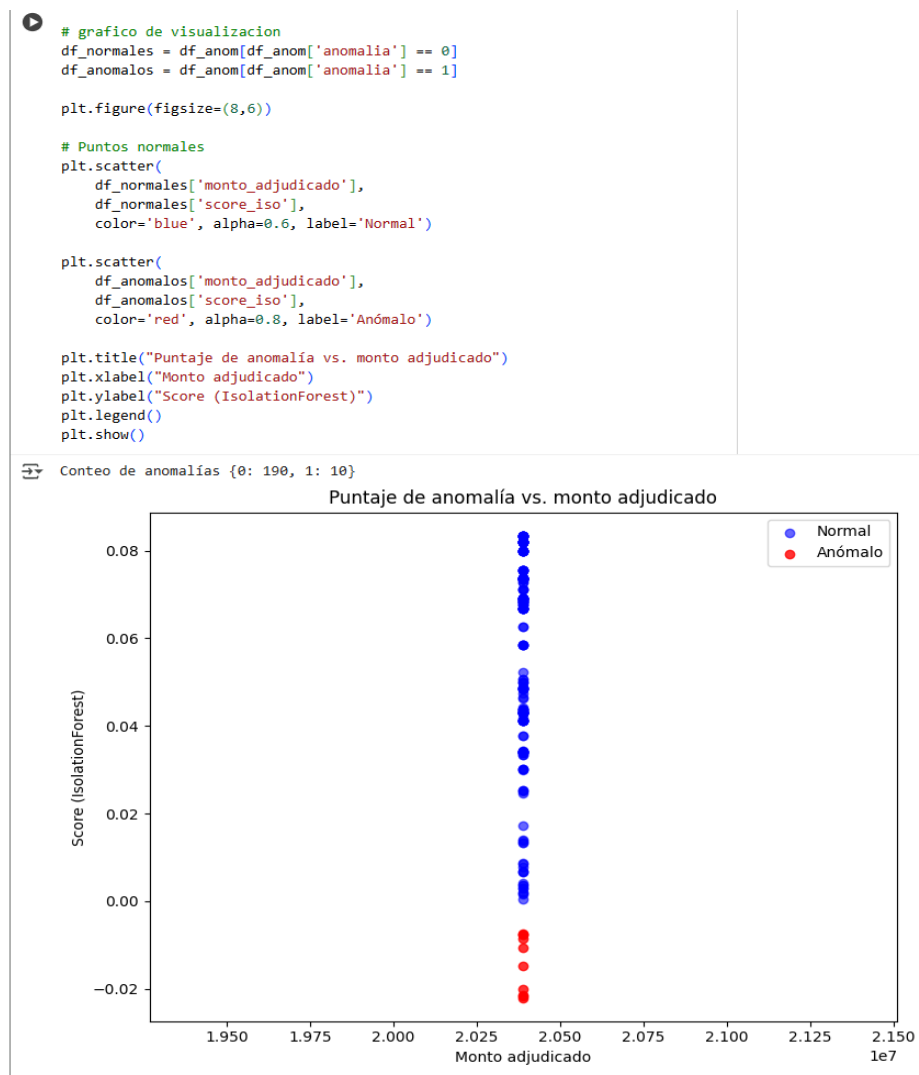
pipe = Pipeline([('pre', pre), ('iso', iso)])
pipe.fit(X)

# Predicciones
X_enc = pipe['pre'].transform(X)
labels = pipe['iso'].predict(X_enc)
scores = pipe['iso'].decision_function(X_enc)

df_anom['anomalía'] = (labels == -1).astype(int)
df_anom['score_iso'] = scores

print("Cuento de anomalías", df_anom['anomalía'].value_counts().to_dict())
```

El desarrollo de este bloque del código implementa una herramienta de auditoría automática. Permitiendo identificar registros que se alejan del comportamiento típico de la contratación, ayudando a detectar posibles errores, irregularidades o sobrecostos, para lograr priorizarlas en procesos contractuales atípicos. Por ejemplo: Si de 1.000 contratos, 50 son marcados como anómalos, el encargado del área puede concentrarse en esos 50 para verificar si hay sobrevaloración, tiempos irregulares o modalidades atípicas.

**Figura 20** Puntaje de anomalía vs. monto adjudicado (Isolation Forest)

Se logra evidenciar la visualización de los resultados del modelo de detección de anomalías mediante el algoritmo Isolation Forest. En este gráfico de dispersión, los puntos azules representan los procesos contractuales normales, mientras que los rojos corresponden a los casos anómalos identificados por el modelo. El eje X muestra el monto adjudicado y el eje Y el puntaje de anomalía (score\_iso), donde los valores más bajos indican comportamientos atípicos.

De acuerdo con la gráfica, se identificaron 10 procesos contractuales anómalos frente a 190 registros normales, lo que equivale aproximadamente al 5 % del total analizado. Estos casos destacan por presentar desviaciones en los patrones esperados de adjudicación y requieren una revisión detallada, ya que podrían corresponder a inconsistencias, errores administrativos o posibles irregularidades en la gestión contractual.

### **Lecciones y recomendaciones**

Durante el desarrollo de este proyecto se logró adquirir un conocimiento fundamental sobre la aplicación de modelos analíticos y predictivos en la gestión de la contratación pública, comprendiendo no sólo los aspectos técnicos del procesamiento y modelamiento de datos, sino también las implicaciones administrativas que conlleva su uso, esta experiencia permitió reconocer la importancia de la calidad de la información, la consistencia de los registros contractuales y la adecuada estructuración de las bases de datos como elementos clave para la construcción de modelos confiables y la generación de pronósticos útiles para la toma de decisiones.

Implementar el modelo Prophet y el algoritmo de detección de anomalías Isolation Forest permitió una visión más clara sobre el comportamiento histórico de los montos adjudicados, la identificación de patrones estacionales y la detección temprana de irregularidades. Este aprendizaje evidenció que la combinación de técnicas de machine learning se convierte en un mecanismo eficaz para realizar un control preventivo, capaz de anticipar riesgos financieros y mejorar la planeación presupuestal. Así mismo, se reafirmó la necesidad de integrar el análisis de datos dentro de las prácticas de gestión pública, no solo como una herramienta técnica, sino como un recurso estratégico para fortalecer la transparencia y la eficiencia administrativa.

Sin embargo, se presentaron algunas limitaciones especialmente relacionadas con la actualización de los datos. Estos factores dificultaron la correcta interpretación de los resultados, limitando la precisión de los modelos predictivos. Por ello, se considera fundamental promover la estandarización de la información contractual, en base a esta experiencia se concluye que el uso de la analítica de datos aplicada a la contratación pública no solo mejora la capacidad de planificación, sino que impulsa una cultura organizacional orientada a la toma de decisiones basada en evidencia y la optimización de recursos.

### Referencias bibliográficas

- Ayobami, A., Mike-Olisa, U., Ogeawuchi, J., Abayomi, A., & Agboola, O. (2023, 01). Algorithmic Integrity: A Predictive Framework for Combating Corruption in Public Procurement through AI and Data Analytics. *Journal of Frontiers in Multidisciplinary Research*, 4(2), 130-141. <https://doi.org/10.54660/.JFMR.2023.4.2.130-141>
- Bowen, D., & Ungar, L. (2020). Generalized SHAP: Generating multiple types of explanations in machine learning. *arXiv [cs.LG]*. <https://doi.org/10.48550/ARXIV.2006.07155>
- Cobo Navarrete, I. (2022). Procedimiento administrativo electrónico. En. *Administración digital*, (51-83). Dykinson
- Gallego, J., Rivero, G. & Martínez, J. (2021). Preventing rather than punishing: An early warning model of malfeasance in public procurement. *International Journal of Forecasting*, 37(1), 360-377. <https://doi.org/10.1016/j.ijforecast.2020.06.006>.
- García Rodríguez, M. J. (2022). Tecnologías digitales para el control de la contratación pública. *Auditoría pública: revista de los Órganos Autónomos de Control Externo*, (79), 88-101. <https://asocex.es/wp-content/uploads/2022/06/5.-Tecnologias-digitales-para-el-control-de-la-contratacion-publica.pdf>
- Genes, D., Zapata Martínez, O. D., & Ortega, M. J. (2021). *Ley 80 de 1993*. Universidad de Córdoba. <https://repositorio.unicordoba.edu.co/handle/ucordoba/4118>
- Górski, J. (2023). Digital transformation of public procurement: WTO's historical and ongoing role. In *The Elgar Companion to the World Trade Organization*, 108-150. <https://doi.org/10.4337/9781800882867.00014>
- Hochstetter, J., Vásquez, F., Diéguez, M., Bustamante, A., & Arango-López, J. (2023). Transparency and E-government in electronic public procurement as sustainable development. *Sustainability*, 15(5), 4672. <https://doi.org/10.3390/su15054672>
- Mishra, S., Shinde, M., Yadav, A., Ayyub, B., & Rao, A. (2024). An AI-Driven Data Mesh Architecture Enhancing Decision-Making in Infrastructure Construction and Public Procurement. *Computer Science*. <https://doi.org/10.48550/arXiv.2412.00224>
- Organisation for Economic Co-operation and Development (OECD). (2021). *Perspectivas económicas de América Latina 2021: Avanzando juntos hacia una mejor recuperación*. OECD Publishing. <https://doi.org/10.1787/2958a75d-es>

- Pérez Astonitas, R. P., Lobatón Arenas, C. L. L., Figueroa Piscocoya, N. F., Taboaba Arana, D. G. T., & Yeckie Arteaga, R. M. Y. (2021). Análisis de datos para la gestión estratégica pública. Data analysis for strategic public management. *Revista Científica Emprendimiento Científico Tecnológico*, (2), 1-11. <https://revista.ectperu.org.pe/index.php/ect/article/view/44>
- Quevedo-Piratova., D.A. , Villalba-Londoño., J.U. y Gonzalez-Gomez. (2024). Application of machine learning for predictions of consecutive dependent data of type  $\{(a, b) \rightarrow c\} \rightarrow d$ . *Tecnura*, 28(79), 66-86. <https://doi.org/10.14483/22487638.22094>
- Ramos-Rivadeneira, D. X., & Jiménez-Toledo, J. A. (2024). La innovación desde las tecnologías emergentes para la competitividad empresarial. *Gestión Y Desarrollo Libre*, 9(17). 1-14. <https://doi.org/10.18041/2539-3669/gestionlibre.17.2024.11052>
- Raschka, S., & Mirjalili, V. (2020). *Python machine learning: aprendizaje automático y aprendizaje profundo con Python, scikit-learn y TensorFlow*. Marcombo.
- Rodríguez Arévalo, S. (2021). Predicción de ineficiencias en la contratación pública de Bogotá. Universidad del Rosario. [https://doi.org/10.48713/10336\\_30915](https://doi.org/10.48713/10336_30915)
- Salazar, A., Pérez, J. F., & Gallego, J. (2024). VigIA: prioritizing public procurement oversight with machine learning models and risk indices. *Data & Policy*, 6, 1-22. <https://doi.org/10.1017/dap.2024.83>
- Siciliani, L., Taccardi, V., Basile, P., Di Ciano, M., & Lops, P. (2023). AI-based decision support system for public procurement. *Information Systems*, 119, 1-16. <https://doi.org/10.1016/j.is.2023.102284>
- Xu, H., Pang, G., Wang, Y., & Wang, Y. (2023). Deep isolation forest for anomaly detection. *IEEE Transactions on Knowledge and Data Engineering*, 35(12), 12591-12604. <https://doi.org/10.1109/TKDE.2023.3270293>