



iStock™
Credit: AmnajKhetsamtip

1449334081

PREDICCIÓN DE PROCESOS JUDICIALES

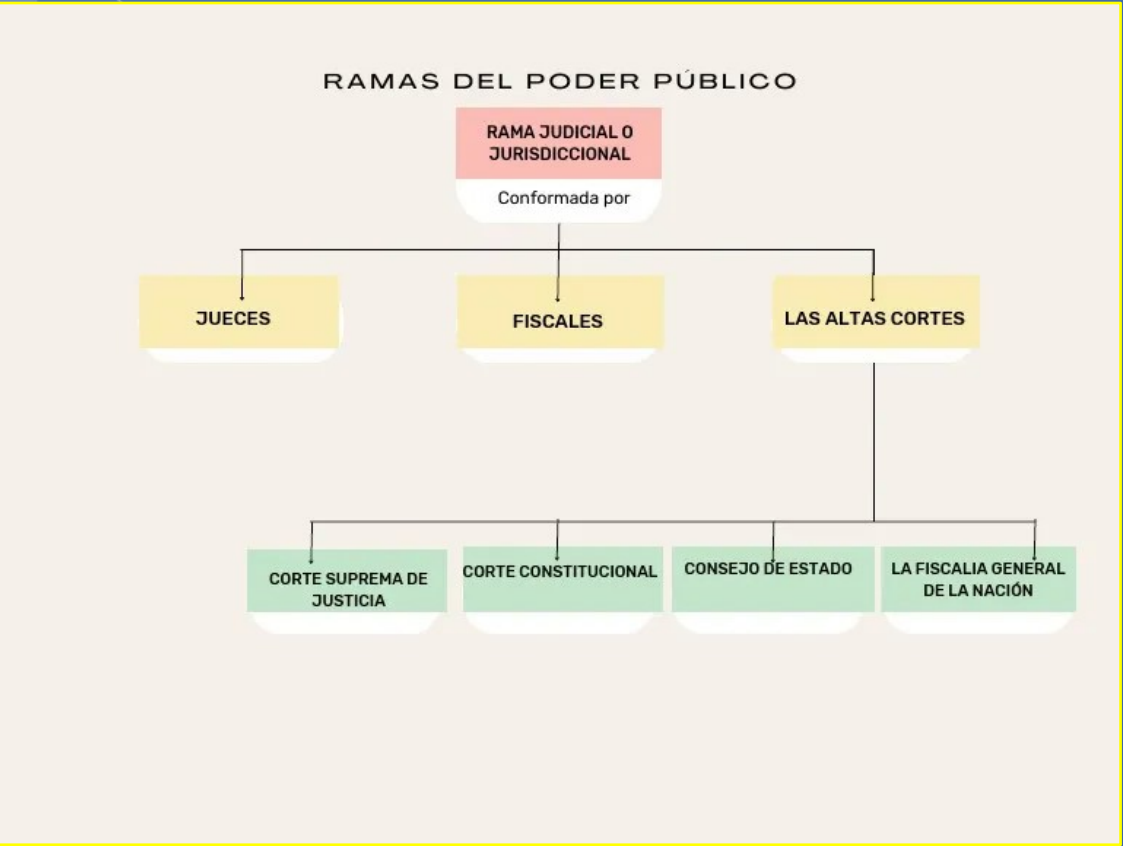
Autor
Roger Acuña - 2025

AGENDA

- Problemática actual
- Objetivo
- Pipeline de ML
- Experimentos

- Deployment
- Conclusiones
- Mejoras - Pasos a seguir

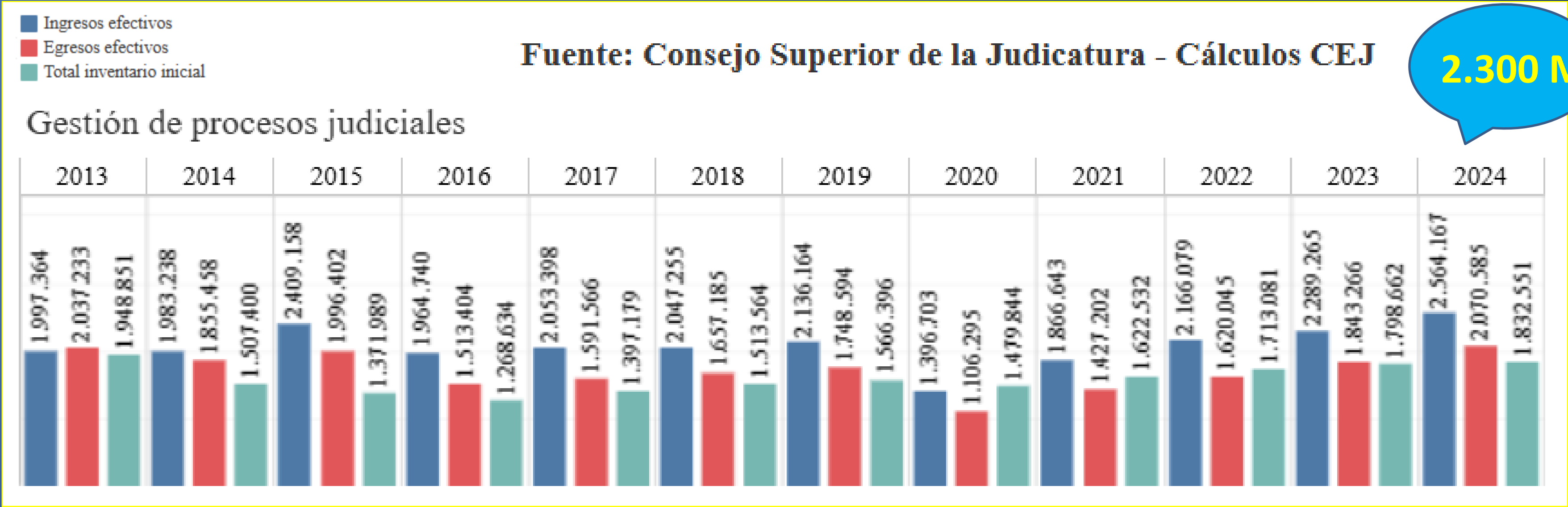




PROBLEMÁTICA ACTUAL RAMA JUDICIAL

INGRESO DE PROCESOS

- No cuenta con información de contexto
- No cuenta con buenos estimativos.
- Se realizan cálculos de pronósticos a corto plazo.

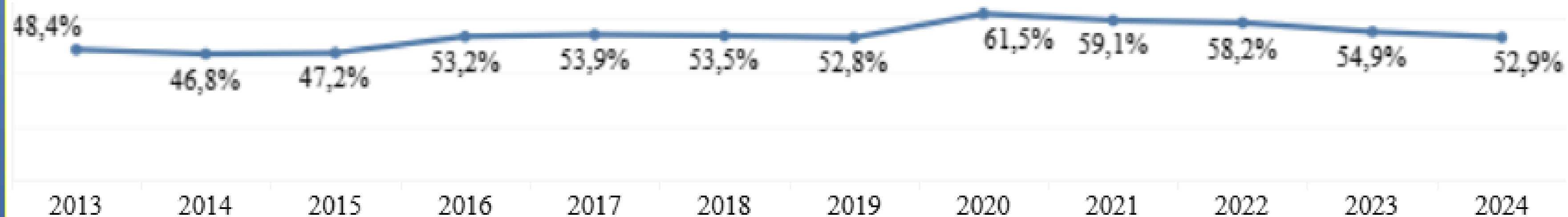


HISTORICOS DE INDICES DE CONGESTION JUDICIAL

Índice de Evacuación Total - IET



Índice de Congestión



Tomado de www.ramajudicial.gov.co

OBJETIVO

1 BOGOTÁ	1 PENAL
2 MEDELLÍN	2 EJEC. PENAS
3 CALI	3 PENAL ADOLESCENTES
4 B/MANGA	4 PENAL ESPECIALIZADO
5 PASTO	5 CONTROL GARANTIAS

Predecir o proyectar la demanda de procesos judiciales en los juzgados penales de 5 ciudades con la finalidad de tomar decisiones asertivas.

La idea es predecir para un período dado (año y trimestre) en una ciudad y especialidad penal, cuantos procesos ingresaran al juzgado.

PIPELINE DE MACHINE LEARNING

Machine Learning Pipeline



RECOLECCIÓN DE LA DATA

- DANE – GEIH
- Fiscalía
- Unidad de Víctimas
- Rama Judicial

fiscalia.gov.co/colombia/gestion/estadisticas/

viernes, 30 de mayo de 2025 10:31 amEnglish SiteTamaño TextoAlto ContrasteConverTICgov.co



EntidadPlaneación y EstrategiaPrensaContrataciónTransparencia y acceso a informaciónAtención y Servicios a CiudadaníaParticipaDenuncia fácil

Inicio > Gestión > Datos abiertos de la Fiscalía General de la Nación

Datos abiertos de la Fiscalía General de la Nación

Estadísticas Gráficas

- Noticia criminal (Agregado de delitos)
- Noticia criminal (Delitos específicos)

Estas estadísticas corresponden a los registros de las notificaciones (SPOA) por presuntos hechos delictivos que la Fiscalía General de la Nación emite en virtud de la vigencia de la Ley 906 de 2004 y la Ley 1098 de 2006 y por el debido a que el sistema de información se implementó a partir de 2010. Los datos se han logrado consolidar desde el año 2010.

Las estadísticas están disponibles de manera gráfica o en formato de archivo CSV para su análisis de comportamiento y la evolución de las entradas de noticias.

Consejo Superior de la JudicaturaCorte Suprema de justiciaConsejo de EstadoCorte ConstitucionalComisión Nacional de la Verdad

5 de Jun - 2025

CONSULTA DE PROCESOS NACIONAL UNIFICADA

Saludos, usted tiene a su disposición la nueva Consulta de Procesos Nacional Unificada (CPNU) en su versión 2.0, cuyo objetivo es entregar a la ciudadanía en general un producto uniforme donde consultar sus procesos.

Es importante señalar que se podrán utilizar temporalmente las dos consultas de procesos existentes junto con la CPNU.

CONSULTA DE PROCESOS NACIONAL UNIFICADA

CONSULTA DE PROCESOS

JUSTICIA XXI WEB

unidadvictimas.gov.co

GOV.CO

Unidad para las Víctimas

INICIO | ATENCIÓN Y SERVICIOS CIUDADANÍA | COMUNICACIONES | LA UNIDAD | TRANSPARENCIA | PARTICIPA

Ingrese su termino de búsqueda aquí...





PORTADA

Avanza cumplimiento de acuerdos con el pueblo Emberá: Unidad para las Víctimas caracterizó a 349 hogares que están en el Parque Nacional

Esta jornada de atención se realizó en articulación con la Alcaldía Mayor de Bogotá, la Defensoría del Pueblo y la Procuraduría, y tuvo el propósito de identificar las necesidades y condiciones de las familias para avanzar en la definición de rutas de retorno, reubicación e integración local.

ANDA | DANE

PÁGINA PRINCIPAL / CATÁLOGO CENTRAL DE DATOS / MERCLAB-MICRODATOS / DANE-DIMPE-GEIH-2024

Gran Encuesta Integrada de Hogares - GEIH - 2024.

Colombia, 2024

Directorio de Metodología y Producción Estadística - DIMPE

Creado el February 12, 2025 Última modificación February 12, 2025 Visitas a la página 358.106 Descargar 156.432 Documentación en PDF

metadatos

Descripción de la operación estadística

Identificación

Alcance

Cobertura

Productores patrocinadores

Muestreo

Recolección de Datos

Cuestionarios

Identificación

Idno

DANE-DIMPE-GEIH-2024

Título

Gran Encuesta Integrada de Hogares - GEIH - 2024.

Título Traducido

ANÁLISIS EXPLORATORIO DE DATOS

12 Variables Predictoras

Variable a Predecir : **TOTAL PROCESOS**

```
[ ] # Conocer en Detalle la estructura del DataSet  
df.info()
```

```
[+] <class 'pandas.core.frame.DataFrame'>  
RangeIndex: 1410 entries, 0 to 1409  
Data columns (total 13 columns):  
#   Column                Non-Null Count  Dtype    
---  ---                  
0   año                   1410 non-null  float64  
1   trimestre             1410 non-null  object   
2   ciudad                1410 non-null  object   
3   subespecialidad       1410 non-null  object   
4   Total_procesos         1410 non-null  float64  
5   TOTAL_INDICIADOS       1410 non-null  float64  
6   TOTAL_NOTICIAS_CRIMINALES 1410 non-null  float64  
7   TOTAL_VICTIMAS_FISCALIA 1410 non-null  float64  
8   VICTIMAS_UNIDAD_VICTIMAS 1410 non-null  float64  
9   Total_Poblacion        1410 non-null  float64  
10  RGN                    1410 non-null  object   
11  COVID                  1410 non-null  object   
12  LEY_1878_2018          1410 non-null  object   
  
dtypes: float64(6), int64(5), object(2)  
memory usage: 143.3+ KB
```

	count	mean
año	1410.0	2017.448809
trimestre	1410.0	2.510638
Total_procesos	1410.0	2328.872340
TOTAL_INDICIADOS	1410.0	3238.712340
TOTAL_NOTICIAS_CRIMINALES	1410.0	4584.889645
TOTAL_VICTIMAS_FISCALIA	1410.0	3718.553282
VICTIMAS_UNIDAD_VICTIMAS	1410.0	87.715603
Total_Poblacion	1410.0	109130.575358
RGN	1410.0	11.944978
COVID	1410.0	0.148938
LEY_1878_2018	1410.0	0.489382

	Cuantos Nulos	Porcentaje
año	0	0.0
trimestre	0	0.0
ciudad	0	0.0
subespecialidad	0	0.0
Total_procesos	0	0.0
TOTAL_INDICIADOS	0	0.0
TOTAL_NOTICIAS_CRIMINALES	0	0.0
TOTAL_VICTIMAS_FISCALIA	0	0.0
VICTIMAS_UNIDAD_VICTIMAS	0	0.0
Total_Poblacion	0	0.0
RGN	0	0.0
COVID	0	0.0
LEY_1878_2018	0	0.0

- 1410 Instancias
- 13 Features
- No tenemos datos nulos

ANALISIS UNIVARIADO

TOTAL PROCESOS

Valores Atípicos

indica valores
atípicos altos.

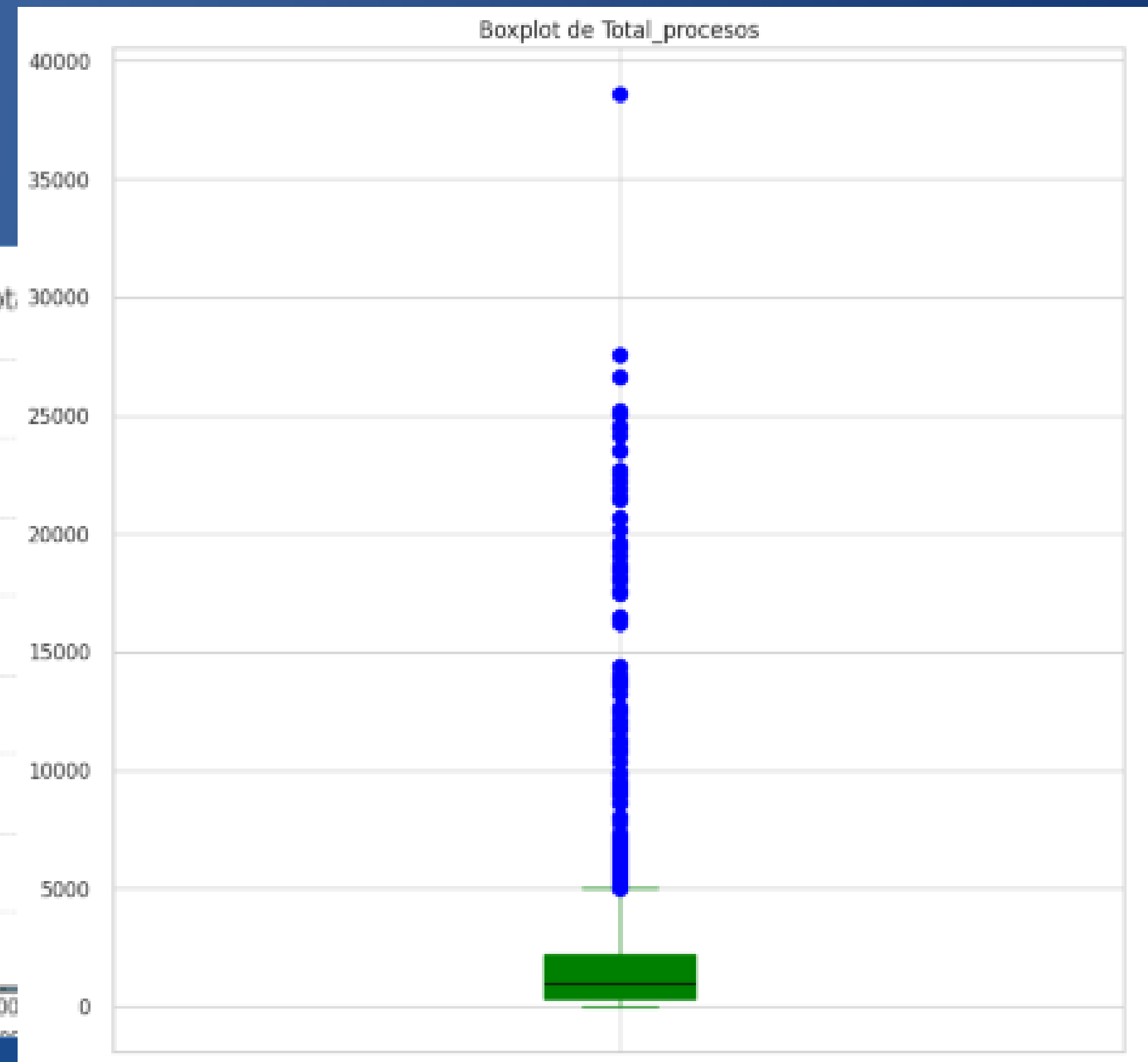
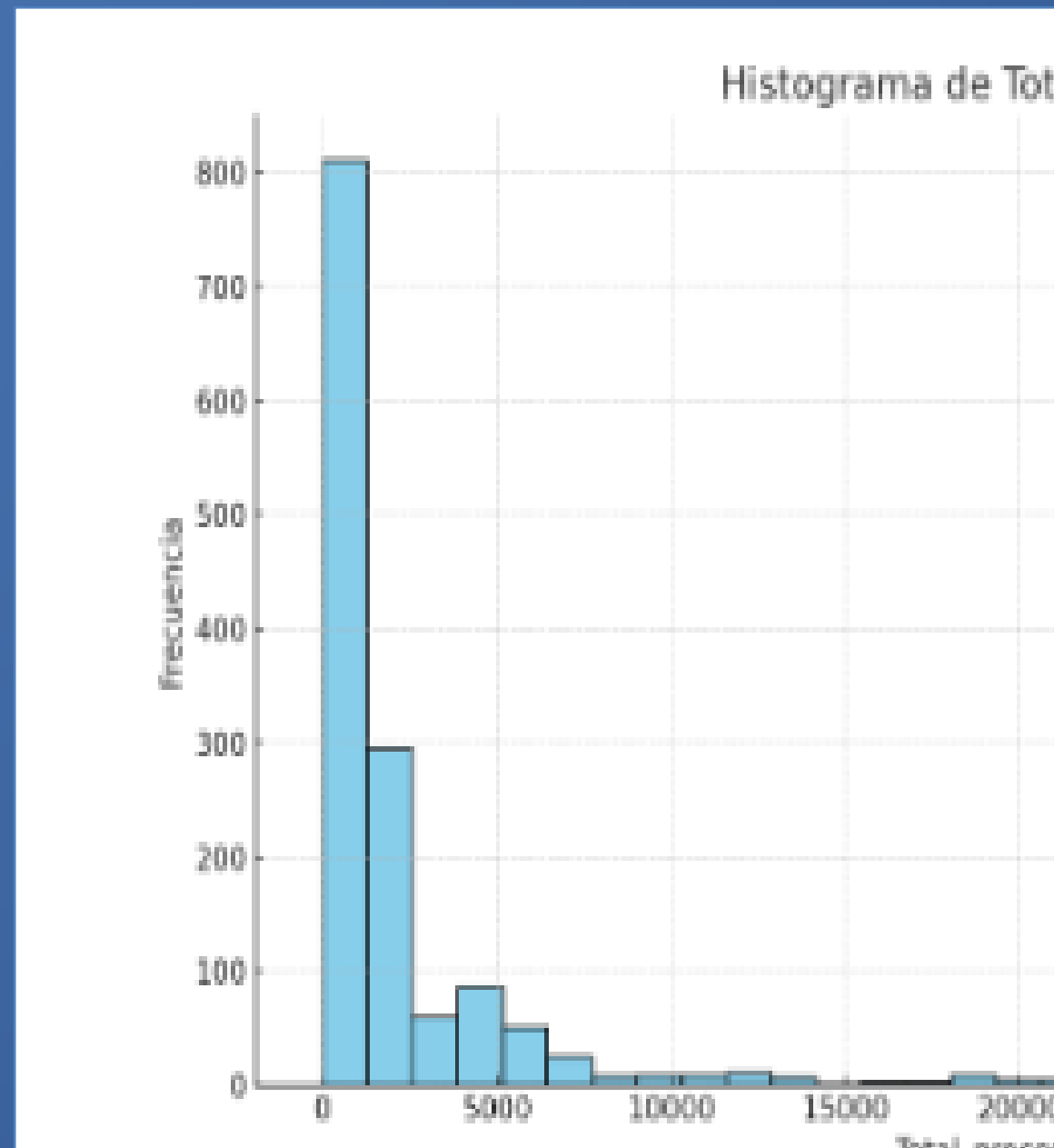
Max valor **(38.596)**

Variabilidad

- 50% tienen menos
de **2.326** procesos.

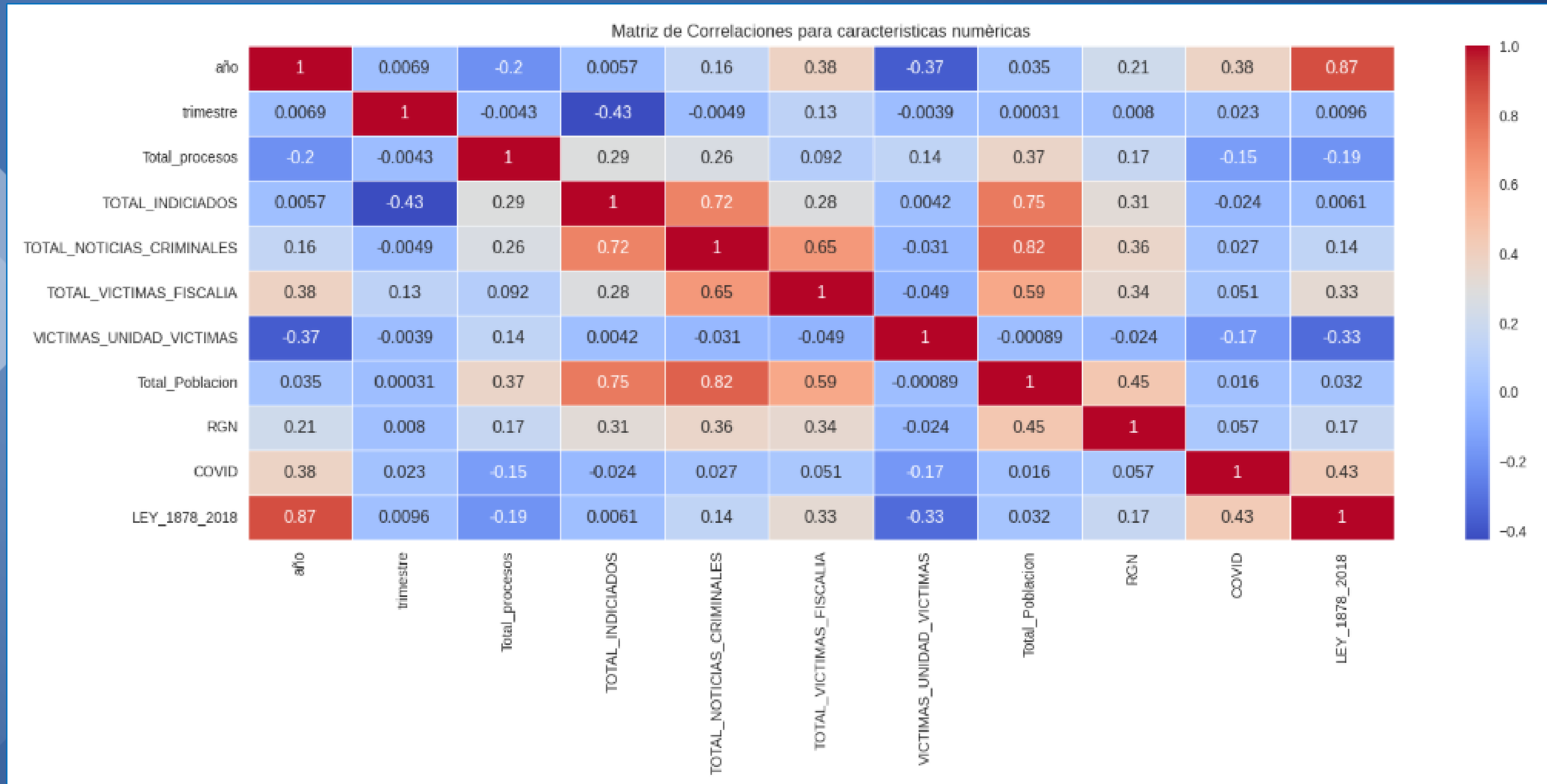
Asimetría

- La distribución es
altamente asimétrica
positiva,
sesgada a la derecha.



		media	Min	Max
Total procesos		2326	0	38596

ANALISIS BIVARIADO – MATRIZ DE CORRELACIONES



Variable dependiente :
Total Procesos

• Total Población (0.37)

• Total Indiciados(0.29)

• Total Noticias
Criminales(0.26)

EXPERIMENTOS

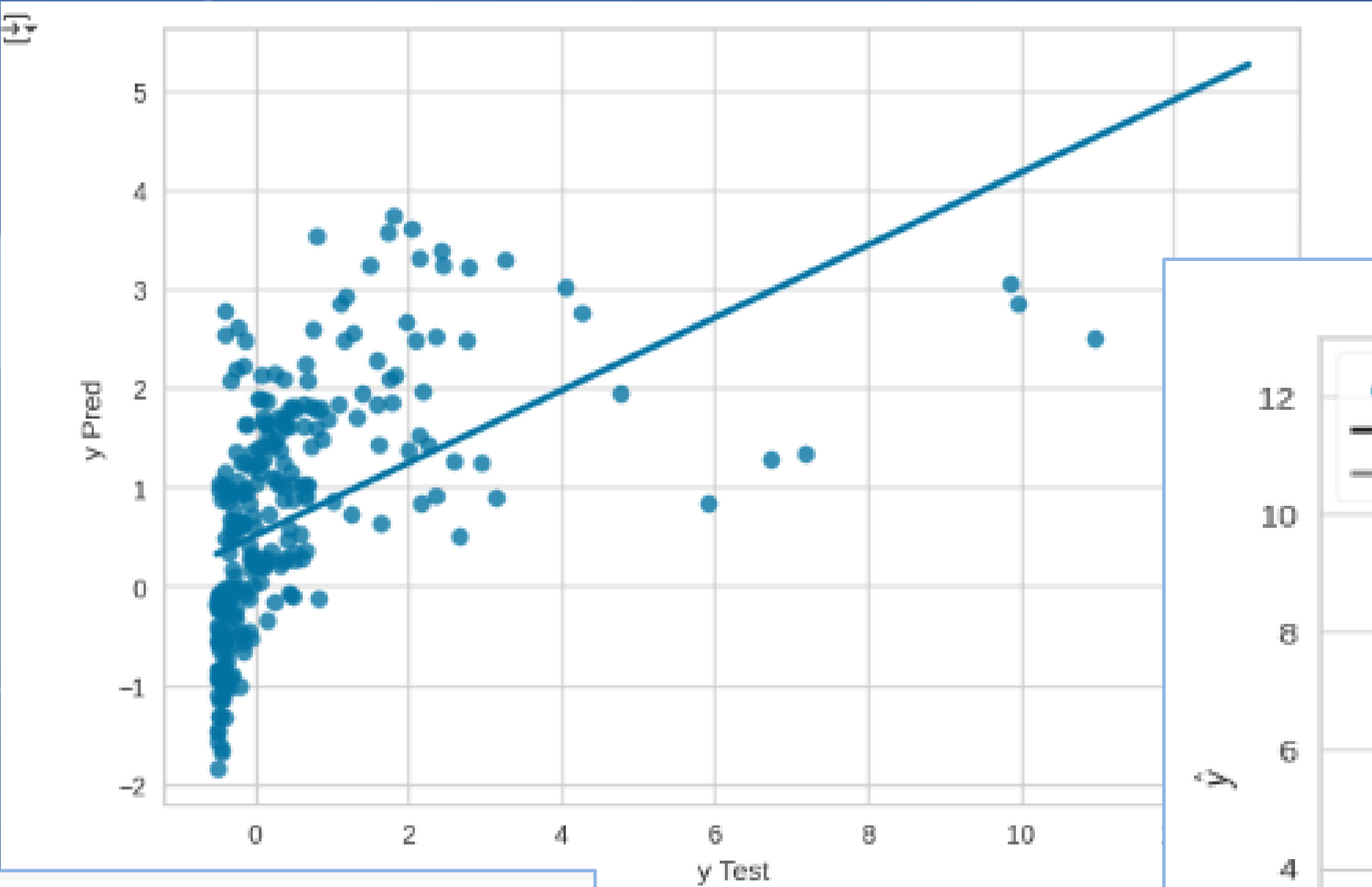


Regresión Lineal

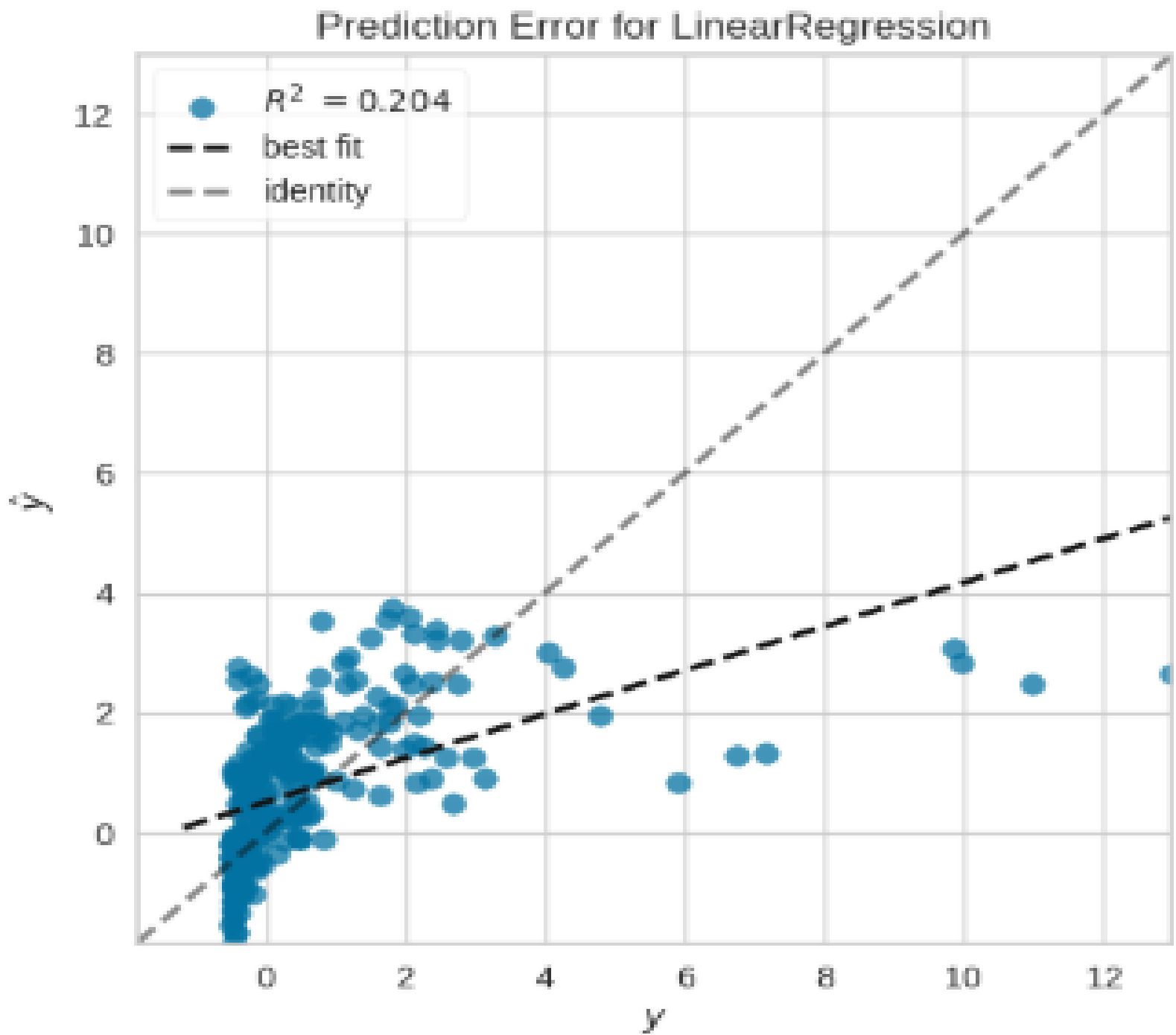
R2 o Accuracy (0 – 1)
R2 = 0.20

```
[ ] # TIP TIP : Fit es el metodo que nos permite  
model_multi_lin_reg.fit(X_train, y_train)
```

```
[ ] LinearRegression  
LinearRegression()
```



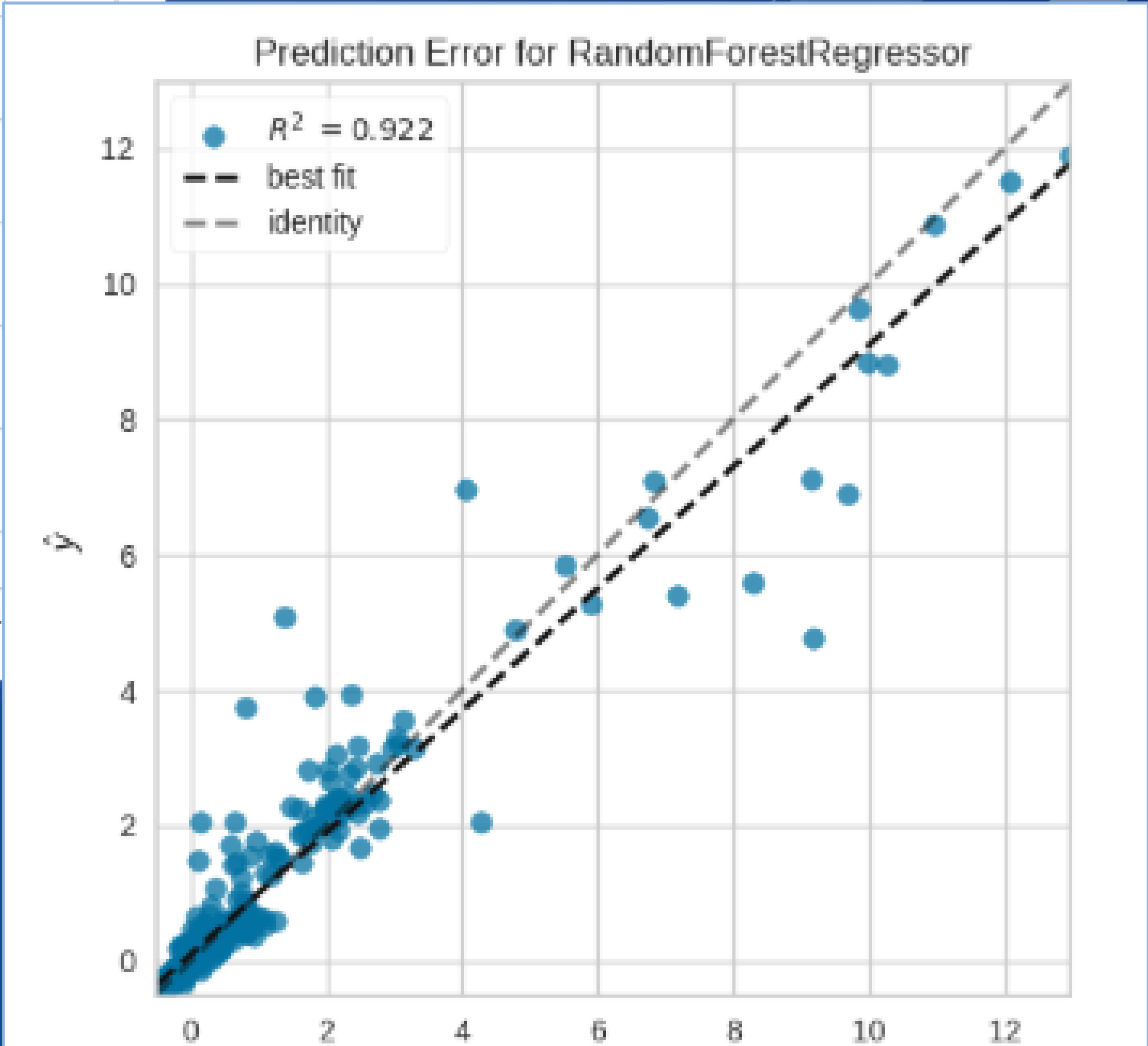
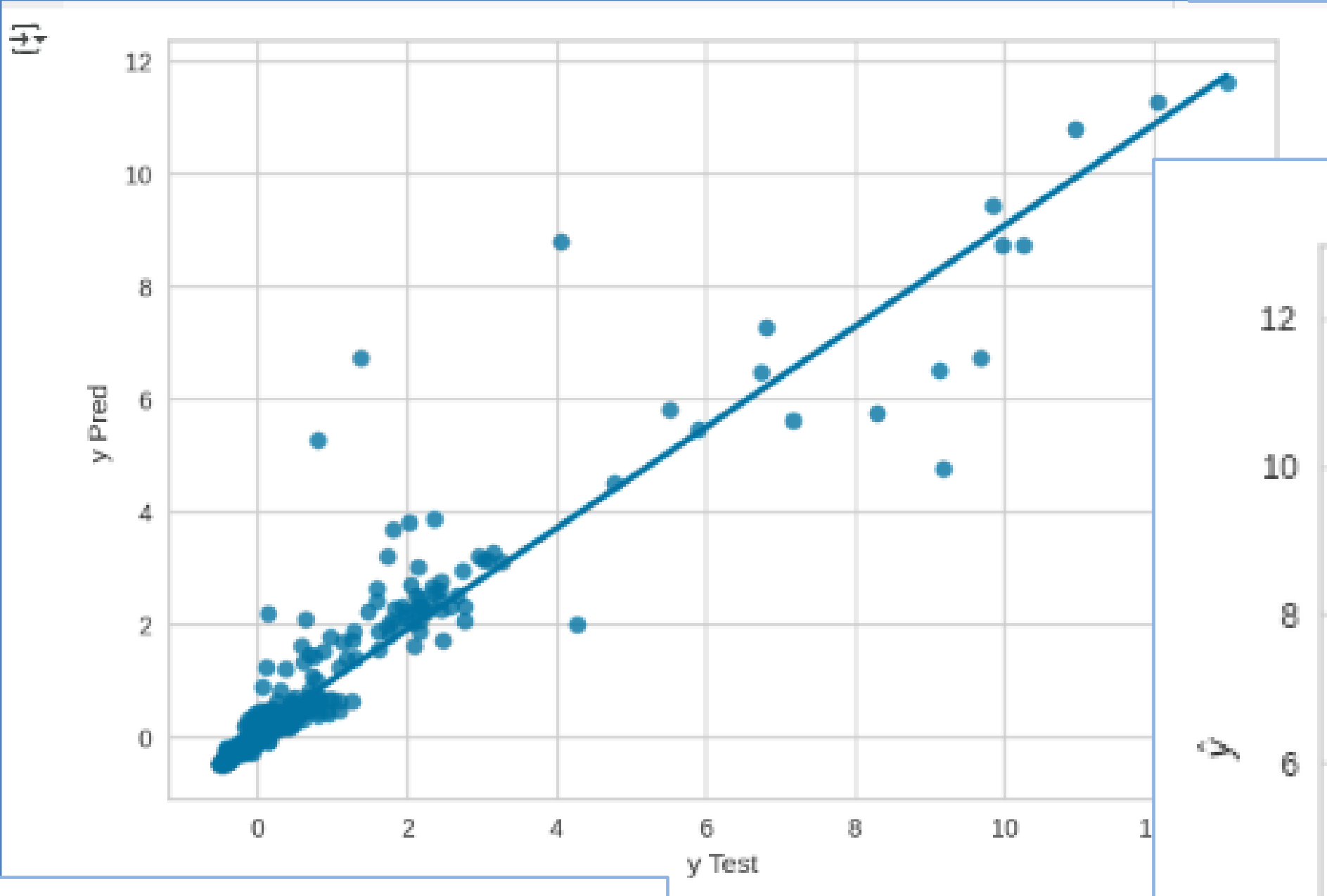
	linear_train	linear_test
R2	0.287827	0.203808
mae	1.072838	0.938354
mse	3.427989	2.285751
rmse	1.851483	1.511870



Random Forest

R2 o Accuracy (0 – 1)
R2 = 0.89

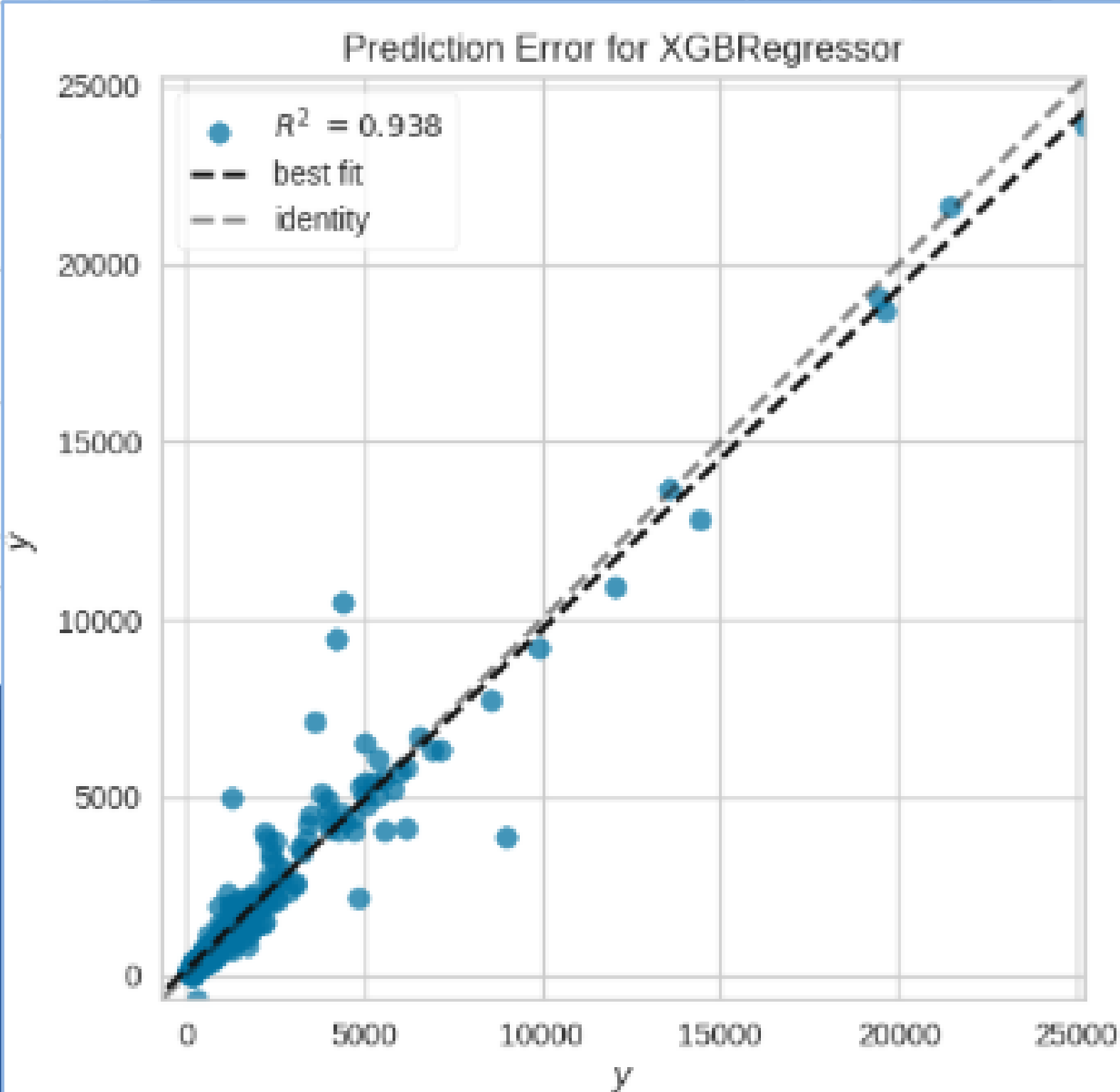
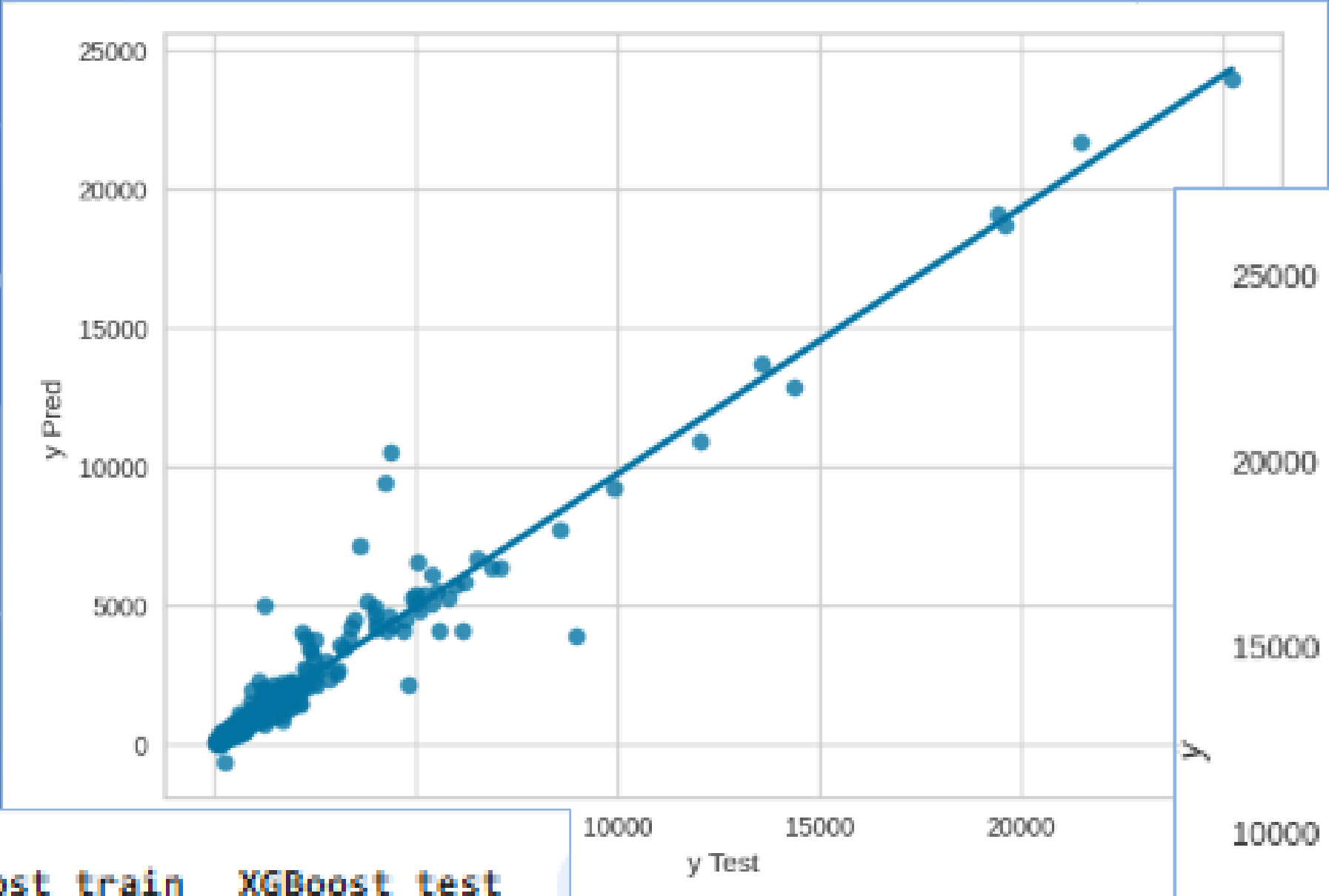
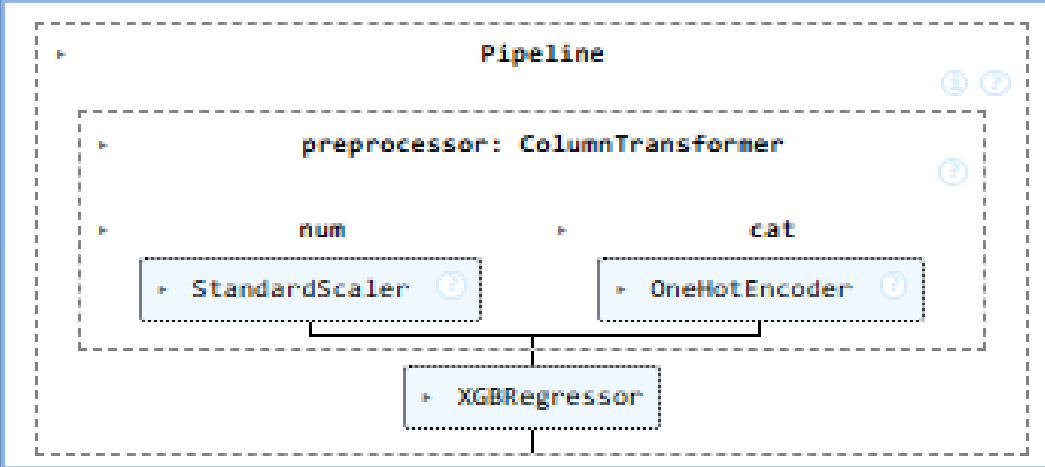
```
RandomForestRegressor
RandomForestRegressor(max_depth=10, min_samples_split=5, random_state=42)
```



	randomforest_train	randomforest_test
R2	0.942008	0.889873
mae	0.144131	0.241381
mse	0.278438	0.392385
rmse	0.527672	0.626407

XGBoostRegression

R2 o Accuracy (0 – 1)
R2 = 0.94



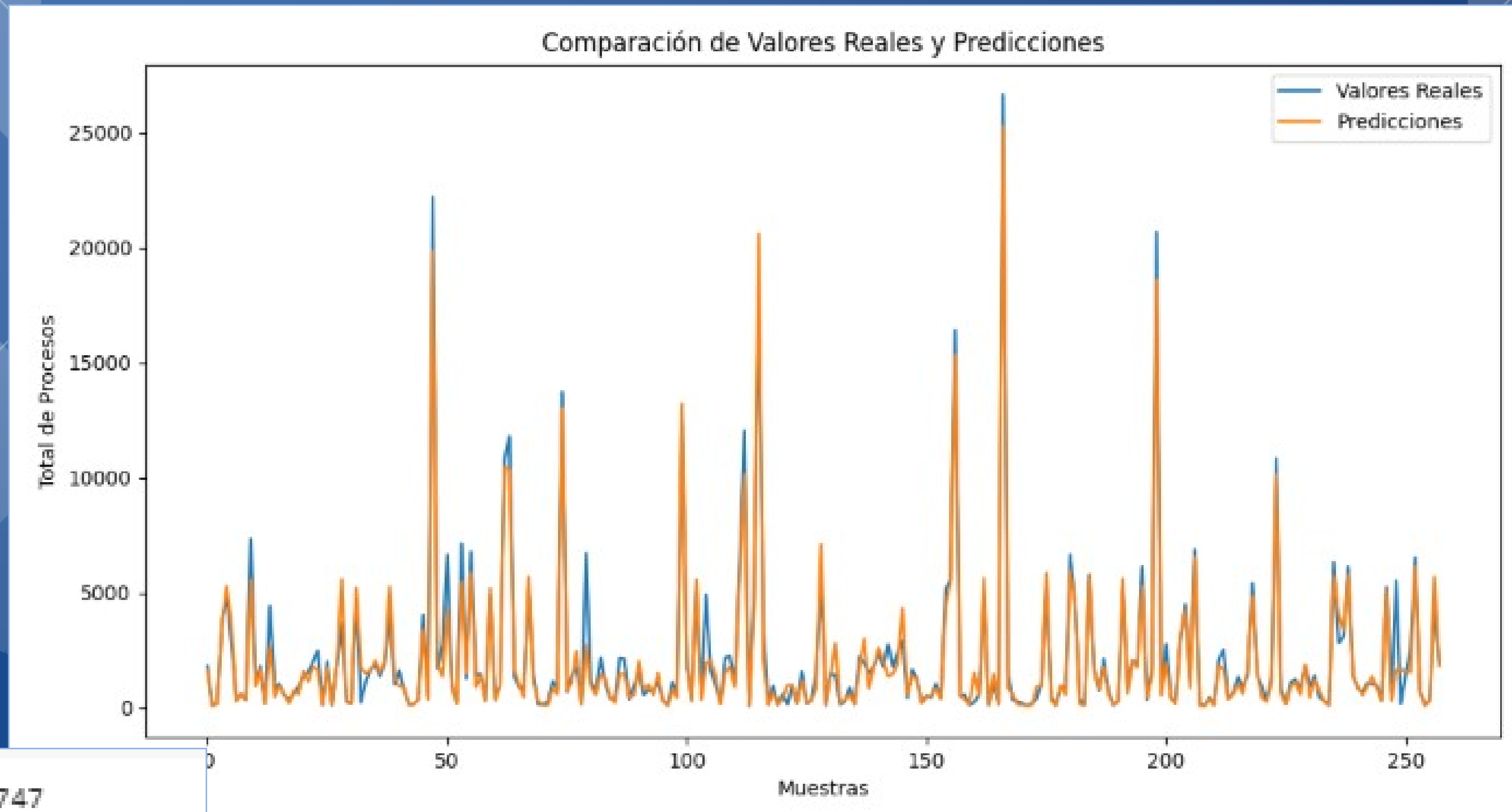
	XGBoost_train	XGBoost_test
R2	0.999982	0.940000
mae	12.814746	350.785248
mse	298.110718	618473.687500
rmse	17.265883	786.430981

LSTM – Redes Neuronales

R2 o Accuracy (0 – 1)

R2 = 0.95

```
Epoch 1/100  
33/33 13s 67ms/step - loss: 0.0557 - val_loss: 0.0389  
Epoch 2/100  
33/33 1s 29ms/step - loss: 0.0466 - val_loss: 0.0358  
Epoch 3/100  
33/33 1s 19ms/step - loss: 0.0398 - val_loss: 0.0358  
Epoch 4/100  
33/33 4s 43ms/step - loss: 0.0376 - val_loss: 0.0333
```



```
MAE: 405.7747  
MSE: 508131.4089  
RMSE: 712.8334  
R²: 0.9594
```



XGBoost Regression

TABLA COMPARATIVA

MODELO	PRECISIÓN	ERROR MEDIO
Regresion lineal	0.2	0.93
Random Forest	0.89	0.24
XGBoost Regression	0.94	350
LSTM - Redes Neuronales	0.95	405

Deployment

Datos de Entrada

Año:

2014

Trimestre:

3

Ciudad:

Pasto

Especialidad:

PENAL PARA ADOLESCENTES CON FUNCION DE CONTROL DE GARANTIAS

Total Población:

16071.25

Total Indiciados:

816.2686867

Total Noticias Criminales:

1405.8

Cobertura Servicios Públicos:

1.776934178

COVID:

0

LEY 1878 DE 2018:

0

Total Unidad Víctimas:

26.875

Total Víctimas Fiscalía:

1247.8

Enviar

Predicción de Procesos Judiciales

```
{
  "Prediccion de Procesos Judiciales": 146
}
```


CONCLUSIONES

01

El modelo XGBoost Regression mostró un buen rendimiento en la predicción de la demanda de procesos judiciales, con métricas de precisión del 0.94 indicando una capacidad razonable para capturar las relaciones entre las variables predictoras y la variable objetivo Total Procesos.

02

El modelo tiene un comportamiento consistente, aunque se identificaron algunos patrones de error como la presencia de valores atípicos o sesgos en ciertos rangos de predicción.

03

El preprocesamiento de datos y toda la ingeniería de features aplicada a los mismos, contribuyó a mejorar el rendimiento del modelo.

OPORTUNIDADES DE MEJORA

PASOS A SEGUIR ...

RE- ENTRENAMIENTO

- Inclusión de mayor cantidad de datos históricos.
- Inclusión de nuevas variables predictoras.
- Eliminar variables no relevantes.
- Inclusión y pruebas de nuevos algoritmos.
- Tratamiento de datos atípicos.
- Generación de datos sintéticos.

APROPIACIÓN DEL MODELO

Apropiación por parte de la Rama Judicial en sus procesos administrativos.

La idea es :

- Mejorar la eficiencia operativa.
- Optimización de recursos.
- Mejorar la satisfacción de los usuarios del Sistema Judicial.

ESCALABILIDAD

- Recopilar resultados reales al futuro y comparar con predicciones.
- Ajustar el modelo en base a errores detectados.
- Mejorar la precisión y disminuir la desviación del error medio absoluto.

The background is a solid dark blue. It features abstract geometric shapes, including triangles and diamonds, in lighter shades of blue, some with thin white lines. At the bottom, there is a silhouette of a city skyline with various skyscrapers.

GITHUB

<https://github.com/rogeracunaherazo/Modelos-Predictivos>

GRACIAS

