

# Análisis Series de Tiempo: Precio del Oro

Nicolás Méndez Gutiérrez - Christian Martinez

2025-05-05



# Contents

<b>1</b>	<b>Descripción del dataset</b>	<b>5</b>
1.1	Estructura de los datos . . . . .	5
<b>2</b>	<b>Objetivo</b>	<b>7</b>
<b>3</b>	<b>Justificación</b>	<b>9</b>
<b>4</b>	<b>Análisis Exploratorio de Datos</b>	<b>11</b>
4.1	Análisis de promedio móvil, rezagos y estacionalidad . . . . .	14
<b>5</b>	<b>Fuentes</b>	<b>21</b>



# Chapter 1

## Descripción del dataset

Este conjunto de datos contiene registros históricos del precio del oro desde el 31 de diciembre de 2013 hasta el 5 de noviembre de 2024, extraídos del mercado MCX. Se trata de un dataset útil para el análisis de series temporales y la predicción de tendencias del precio del oro.

### 1.1 Estructura de los datos

El dataset incluye 2806 entradas y múltiples columnas con información clave:

- Date: Fecha de transacción.
- Open: Precio de apertura del mercado
- High: Precio más alto alcanzado en el día
- Low: Precio más bajo alcanzado en el día
- Price: Precio de cierre del día
- Volume: Cantidad de transacciones realizadas
- Chg%: Variación porcentual del precio respecto al día anterior.



## Chapter 2

# Objetivo

Con este dataset se busca analizar el comportamiento histórico de los precios del oro con el objetivo de identificar tendencias significativas a lo largo del tiempo. Además, permite aplicar modelos estadísticos y de aprendizaje automático para predecir movimientos futuros del mercado, proporcionando herramientas útiles para comprender las dinámicas de este activo financiero y apoyar la toma de decisiones en contextos económicos y de inversión.





## Chapter 3

# Justificación

El dataset de precios diarios del oro es altamente adecuado para la aplicación de técnicas de análisis de series de tiempo debido a las siguientes características:

1. Datos de tiempo ordenados
  - La variable Date proporciona una secuencia cronológica continua sin valores faltantes, lo que facilita la modelación de tendencias y patrones temporales.
  - La periodicidad diaria permite estudiar el comportamiento del mercado con alta resolución temporal.
2. Evolución de variable dependiente
  - La columna Price representa el precio de cierre, una métrica clave para analizar la evolución del valor del oro a lo largo del tiempo.
  - Al ser una serie numérica con cambios graduales y picos específicos, es ideal para aplicar modelos predictivos como ARIMA, modelos de suavizamiento exponencial y redes neuronales recurrentes.
3. Factores Exógenos y Multivariabilidad
  - Las variables Open, High, Low y Volume permiten estudiar la influencia de distintos factores sobre la variación del precio, enriqueciendo el análisis.
  - La columna Chg% proporciona información sobre volatilidad y puede usarse para identificar momentos de alta inestabilidad en el mercado.
4. Aplicabilidad Real y Relevancia Económica
  - El oro es un activo financiero de gran importancia en la economía global, por lo que analizar sus precios a lo largo del tiempo tiene aplicaciones prácticas en predicción de tendencias, evaluación de riesgos y toma de decisiones de inversión.

- Permite la identificación de patrones estacionales, ciclos de mercado y efectos de eventos económicos en la fluctuación del precio.

5. Calidad de los datos

- Con 2806 registros sin valores faltantes, el dataset proporciona información confiable para entrenar modelos sin la necesidad de una limpieza exhaustiva.
- La estabilidad en la estructura del dataset facilita la aplicación de metodologías estadísticas y de machine learning.

En conclusión este dataset es ideal para estudios de series de tiempo, ya que permite aplicar modelos predictivos, evaluar la influencia de factores exógenos y analizar tendencias económicas con datos sólidos y completos.

## Chapter 4

# Análisis Exploratorio de Datos

Se importa el dataset cuyos primeros registros se muestran a continuación.

```
# A tibble: 6 x 7
  Date      Price Open  High   Low Volume `Chg%`
<date>    <dbl> <dbl> <dbl> <dbl> <dbl>   <dbl>
1 2024-11-06 77030 78300 78570 77030     0  -1.86
2 2024-11-05 78490 78224 78670 78160     0   0.11
3 2024-11-04 78401 78498 78642 78237     0  -0.54
4 2024-11-01 78829 78650 78887 78550     0   0.64
5 2024-10-31 78326 79264 79999 77803    90  -1.17
6 2024-10-30 79257 79119 79375 78888   130   0.5
```

A continuación se presenta un resumen de medidas descriptivas.

**Variable type: Date**

Table 4.1: Data summary

Name	datos
Number of rows	2806
Number of columns	7
Column type frequency:	
Date	1
numeric	6
Group variables	
None	

skim_variable	n_missing	complete_rate	min	max	median	n_unique
Date	0	1	2014-01-01	2024-11-06	2019-05-27	2806

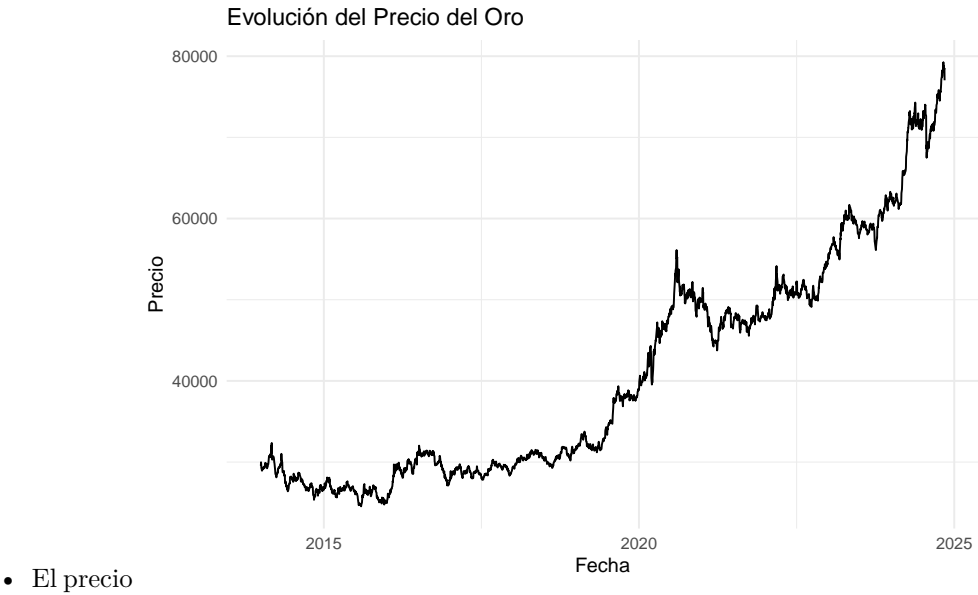
Variable type: numeric

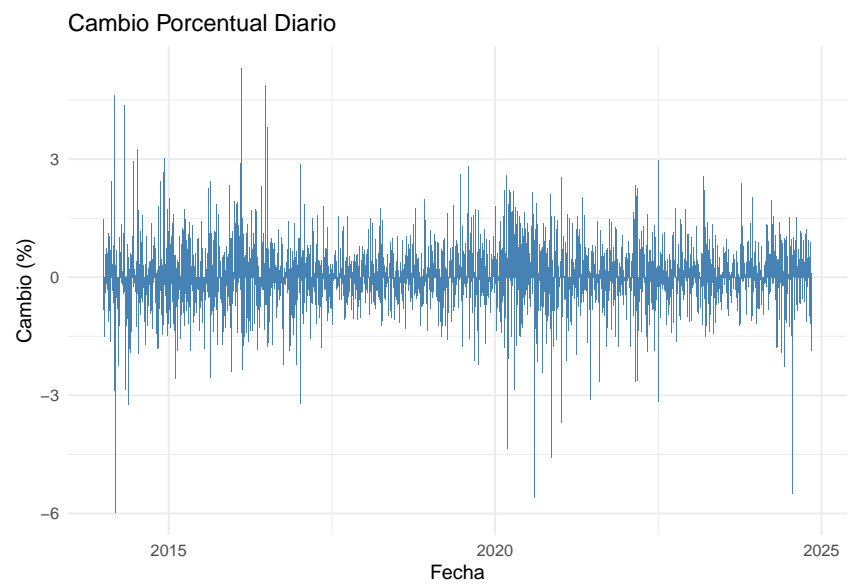
skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50
Price	0	1	40699.89	13828.62	24545.00	29128.00	32980.00
Open	0	1	40700.22	13826.94	24583.00	29103.75	33000.00
High	0	1	40917.78	13900.47	24635.00	29261.25	33220.50
Low	0	1	40482.31	13756.09	24470.00	28974.00	32890.00
Volume	0	1	12529.58	10649.99	0.00	6282.50	10770.00
Chg%	0	1	0.04	0.83	-5.98	-0.38	0.04

Se explora la existencia de datos faltantes.

Date	Price	Open	High	Low	Volume	Chg%
0	0	0	0	0	0	0

Se realizan gráficos para observar tendencias a lo largo del tiempo de:





- Cambios porcentuales



- Comparación del valor máximo vs mínimos del día

## 4.1 Análisis de promedio móvil, rezagos y estacionalidad

### 4.1.1 Promedio móvil

Se agrega un promedio móvil de 7 días, es decir, de manera semanal ya que muchos mercados (como el oro, acciones, productos básicos) tienden a mostrar variaciones semanales (por factores como fin de semana, cierres de mercado, ciclos de noticias, etc.)

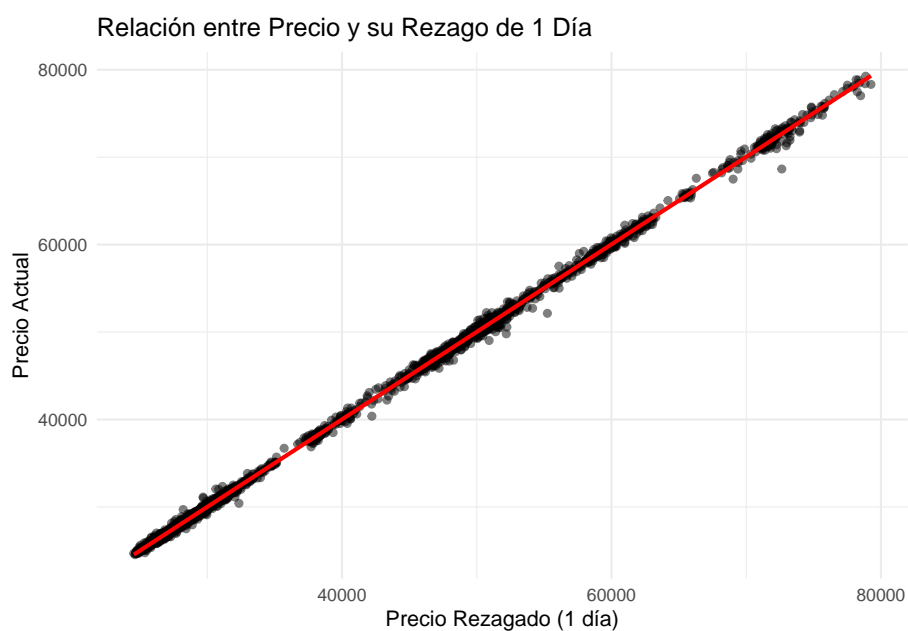


Se hacen las siguientes observaciones:

- La serie muestra una tendencia creciente de largo plazo, especialmente desde 2019 en adelante.
- 2013 - 2018: El precio del oro estuvo relativamente estable o ligeramente a la baja, con pequeñas fluctuaciones.
- 2019 - 2020: Se observa un fuerte crecimiento, con un aumento pronunciado en el precio.
- 2020 - 2021: Hay una corrección o caída parcial, después de un máximo.
- 2021 - 2025: Retoma una tendencia alcista constante con algunos ciclos de subida y bajada.
- Se identifican momentos donde la curva cambia de pendiente (subidas abruptas o correcciones), que pueden estar asociadas a eventos macroeconómicos.

#### 4.1. ANÁLISIS DE PROMEDIO MOVIL, REZAGOS Y ESTACIONALIDAD<sup>15</sup>

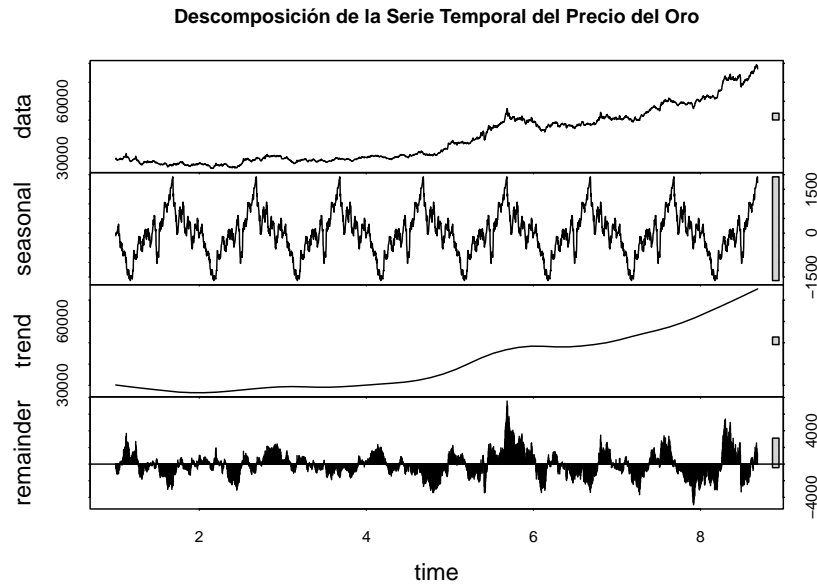
##### 4.1.2 Rezagos (lags)



Se realizan las siguientes observaciones:

- El precio del oro no cambia drásticamente de un día para otro; más bien tiende a seguir la misma trayectoria.
- Hay baja volatilidad diaria relativa (aunque a largo plazo se observaron tendencias importantes).

### 4.1.3 Estacionalidad



- En la primera porción de la gráfica, se observa el comportamiento de la variable precio a lo largo del tiempo.
- La segunda porción, Seasonal, muestra cómo varía sistemáticamente a lo largo del año.
- La tercera porción, muestra la Tendencia a largo plazo.
- Finalmente, se observa en la porción de Remainder el ruido no se explicado por la estacionalidad.

De la gráfica, se realizan las siguientes observaciones:

- Se aprecia claramente la tendencia creciente fuerte, especialmente desde el año 2020 en adelante.
- El componente estacional muestra ciclos repetitivos con una frecuencia regular de picos y valles aproximadamente cada año. La amplitud del patrón estacional es pequeña en comparación al nivel del precio.

Se verifica la estacionalidad.

#### Augmented Dickey-Fuller Test

```
data: serie_ts
Dickey-Fuller = -1.2393, Lag order = 14, p-value = 0.9002
alternative hypothesis: stationary
```

- El ADF test evalúa la hipótesis nula de que una serie no es estacionaria

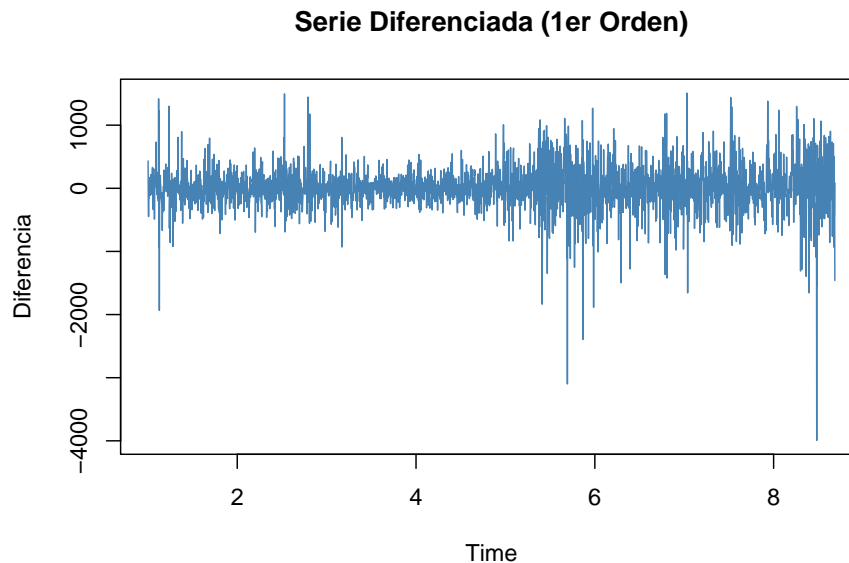


#### 4.1. ANÁLISIS DE PROMEDIO MOVIL, REZAGOS Y ESTACIONALIDAD<sup>17</sup>

(tiene raíz unitaria), frente a la hipótesis alternativa de que sí es estacionaria.

- Con un p-valor tan alto (0.9002), la evidencia estadística indica que la serie `ts_price`:
  - No tiene media ni varianza constante en el tiempo.
  - Tiene una tendencia persistente, como ya se vio en la descomposición STL.
  - No es apta para modelado directo con ARIMA u otras técnicas que requieren estacionariedad, a menos que se transforme.

##### 4.1.4 Diferenciación de primer orden



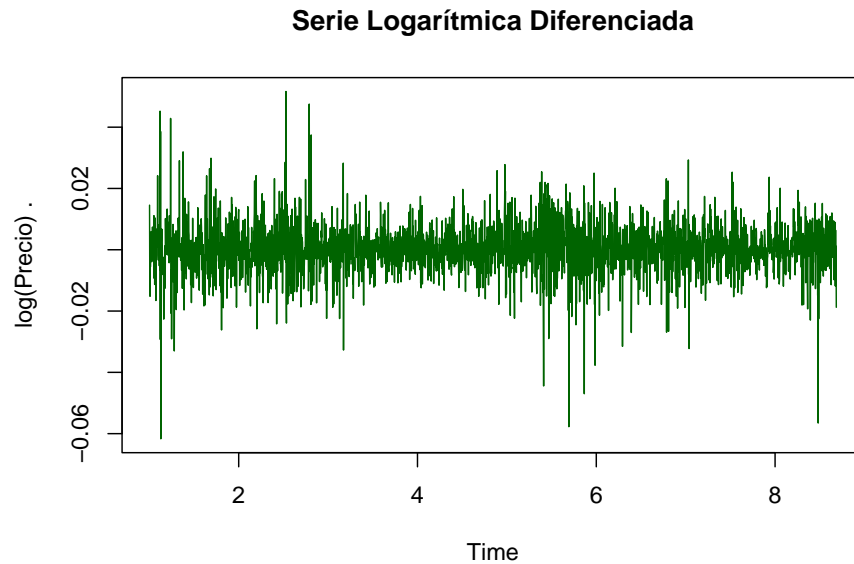
Prueba ADF tras la diferenciación:

##### Augmented Dickey-Fuller Test

```
data: ts_diff1
Dickey-Fuller = -15.197, Lag order = 14, p-value = 0.01
alternative hypothesis: stationary
```

- p-value = 0.01: Significa que podemos rechazar la hipótesis nula de no estacionariedad.
- El estadístico Dickey-Fuller altamente negativo (-15.197) indica una fuerte evidencia de estacionariedad.

### 4.1.5 Transformación logarítmica + diferencia



- La serie resultante oscila en torno a cero, lo cual indica estacionariedad en media.
- No hay una tendencia visible: el promedio es constante.
- La dispersión (volatilidad) se ve bastante estable a lo largo del tiempo  $\rightarrow$  varianza constante, o al menos más homogénea que la serie original.
- Hay algunos picos puntuales que pueden ser eventos de mercado extremos, pero no afectan la estructura general.
- La serie logarítmica diferenciada cumple con los requisitos clave de una serie estacionaria: media constante, varianza relativamente constante y ausencia de tendencia.
- Es altamente recomendable trabajar con esta serie transformada para fines de modelado, predicción o análisis estadístico.

ADF para log-diff:

#### Augmented Dickey-Fuller Test

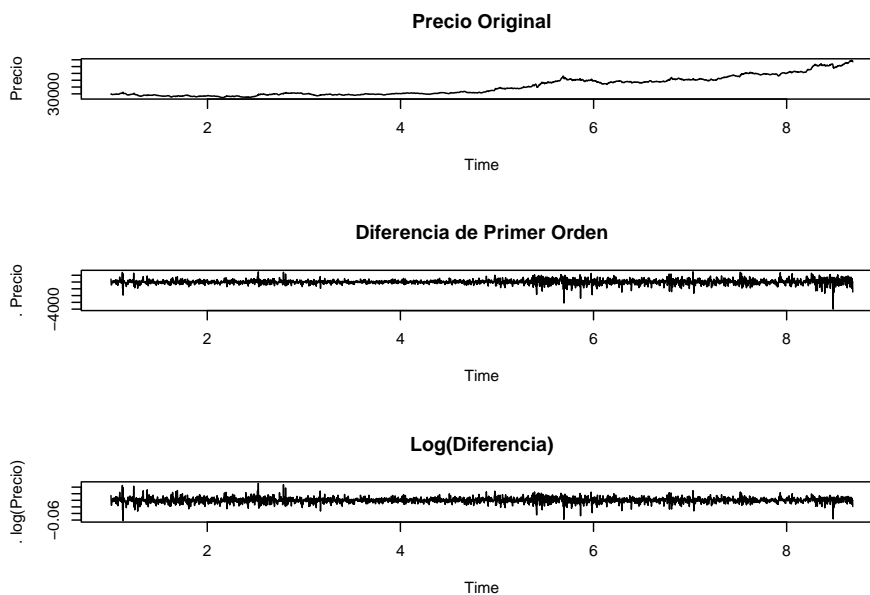
```
data: ts_log_diff
Dickey-Fuller = -14.322, Lag order = 14, p-value = 0.01
alternative hypothesis: stationary
```

- Se aplicó la transformación logarítmica y diferenciación de primer orden a la serie del precio del oro con el objetivo de controlar la tendencia creciente y la variabilidad no constante observadas en la serie original.

#### 4.1. ANÁLISIS DE PROMEDIO MOVIL, REZAGOS Y ESTACIONALIDAD19

- La prueba de Dickey-Fuller aumentada aplicada a la serie transformada ( $\text{diff}(\log(\text{Precio}))$ ) arrojó un estadístico de -14.322 y un p-valor inferior a 0.01, lo cual permite rechazar la hipótesis nula de no estacionariedad.
- Por tanto, se concluye que la serie logarítmica diferenciada es estacionaria, lo cual justifica su uso en procesos de modelado como ARIMA, SARIMA o técnicas de pronóstico más avanzadas. Esta transformación también normaliza la escala de los cambios, permitiendo interpretar los resultados en términos de retornos porcentuales diarios.

##### 4.1.6 Visualización comparativa



- La serie original del precio del oro presenta una tendencia creciente y varianza heterogénea, por lo que no cumple los requisitos de estacionariedad
- La diferenciación de primer orden elimina la tendencia, estabilizando la media pero no completamente la varianza.
- Al aplicar una transformación logarítmica seguida de una diferencia, se consigue una serie que oscila alrededor de cero y mantiene varianza aproximadamente constante.



## Chapter 5

### Fuentes

- Dataset: <https://www.kaggle.com/datasets/nisargchodavadiya/daily-gold-price-20152021-time-series>