**Universidad Tecnológica de Panamá**

Maestría en Analítica de Datos

**Curso:**

Modelos Predictivos

**Avance de Proyecto Final**

**Análisis de precio de la onza de oro, estudio desde 2014 al 2024**

**Profesor:**

Juan Marcos Castillo, PhD

**Elaborado por:**

Edison Reyes Cedeño, 2-708-740

2025

## Introducción

El presente informe tiene como objetivo principal analizar y comprender la evolución histórica del precio de la onza de oro desde enero de 2014 hasta diciembre de 2024. Para ello, se abordará un estudio detallado de esta serie temporal, con el fin de identificar patrones, tendencias y fluctuaciones que han caracterizado este mercado a lo largo de una década. Además de su valor intrínseco, el oro desempeña un papel fundamental como inversión refugio en períodos de incertidumbre económica, convirtiéndose en una opción clave para proteger el capital frente a la volatilidad de los mercados financieros.

Como parte del análisis, se procederá a realizar una predicción a corto y mediano plazo empleando modelos de series temporales: ARIMA, SARIMA y Winter. Concretamente se van a comparar estos 3 modelos utilizando herramientas en Python para alcanzar este cometido.

## Justificación

El oro ha sido históricamente un refugio de valor en tiempos de incertidumbre económica. Es también por tal motivo que se quiere ver cómo impactaron los eventos geopolíticos, la pandemia y crisis financieras a lo largo de esos años y que tan bien puede adaptarse el modelo para previsiones futuras.

Un análisis del comportamiento del precio del oro nos proporciona elementos para:

* Entender patrones históricos y su relación con los mercados financieros
* Modelar tendencias y estacionalidades que ayudan a prever su comportamiento hacia futuro
* Servir de guía en la toma de decisiones relacionadas con inversiones e incluso de política monetaria.

## Antecedentes

Los modelos ARIMA y SARIMA han sido utilizados ampliamente para identificar patrones lineales de series temporales financieras, incluyendo metales preciosos.

El modelo Holt-Winters (Triple Exponential Smoothing) ha sido útil en contextos donde existen tendencias y estacionalidad, como en las investigaciones de Atsalakis y Valavanis (2009) aplicadas a materias primas.

Últimamente se aplican modelos como Prophet (Facebook) o redes neuronales (LSTM) principalmente para captar no linealidades o efectos de eventos externos.

Este trabajo busca realizar una comparación de los modelos ARIMA, SARIMA y Winter para ver qué modelo se adapta mejor y así poder realizar predicciones a futuro.

## Definición del problema

Con los datos históricos de los precios de la onza de oro entre enero de 2014 y diciembre de 2024 se aplicará los modelos de ARIMA, SARIMA y Winter con la ayuda de herramientas disponible en lenguaje Python.

Se pretende evaluar y comparar el desempeño de los modelos e identificar patrones de tendencia y estacionalidad, así como también evaluarlos según las métricas de error.

Se realizará una proyección a 60 días y se evaluará el rendimiento con un conjunto de datos basados en datos trimestrales con el objetivo de también extender la proyección a un mediano plazo, como por ejemplo 5 años.

## Avance de análisis predictivo

Se realiza un breve análisis exploratorio EDA con el fin de obtener una idea general de los datos a ser manipulados.

Estos datos fueron extraídos de la plataforma financiera <https://www.investing.com/> y los mismos muestran una consistencia en el periodo estudiado.

### **Números de registros y variables:**

### Dataframe original

df\_original.info()

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 2816 entries, 0 to 2815

Data columns (total 7 columns):

# Column Non-Null Count Dtype

--- ------ -------------- -----

0 Date 2816 non-null object

1 Price 2816 non-null object

2 Open 2816 non-null object

3 High 2816 non-null object

4 Low 2816 non-null object

5 Vol. 2814 non-null object

6 Change % 2816 non-null object

dtypes: object(7)

memory usage: 154.1+ KB

### **Índice**

Para trabajar con las series de tiempo en lugar de utilizar el índice numérico y consecutivo del dataset, se realizó el cambio tomando el índice como la fecha y la variable Price se transformó a float con el objetivo de realizar los cálculos.

<class 'pandas.core.frame.DataFrame'>

DatetimeIndex: 2816 entries, 2024-12-31 to 2014-01-02

Data columns (total 1 columns):

# Column Non-Null Count Dtype

--- ------ -------------- -----

0 Price 2816 non-null float64

dtypes: float64(1)

memory usage: 44.0 KB

**Valores nulos**

gold\_price\_df.isnull().sum()

|  | **0** |
| --- | --- |
| **Price** | 0 |

**dtype:** int64

No se cuentan con valores nulos en esta serie de tiempo.

### **Estadísticas descriptivas**

gold\_price\_df.describe()

|  | **Price** |
| --- | --- |
| **count** | 2816.000000 |
| **mean** | 1576.401669 |
| **std** | 387.842859 |
| **min** | 1049.600000 |
| **25%** | 1256.750000 |
| **50%** | 1417.550000 |
| **75%** | 1850.275000 |
| **max** | 2800.800000 |

Aquí podemos distinguir la cantidad de registros, el promedio en USD $1576.40, el valor mínimo ubicándose en los USD $1049.60 y el máximo en USD $2800.80

## Gráficas de visualización

**Distribución de frecuencias**

Se realizó un análisis de distribución de frecuencia para ver en qué intervalos se agrupan los precios del oro.

A graph of a number of blue bars

Description automatically generated with medium confidence

**Gráfico histórico desde enero de 2014 a diciembre de 2024**

A green line graph with numbers

Description automatically generated

**Análisis de componentes de la serie de tiempo:**

A graph of a price

Description automatically generated with medium confidence