

# ACTIVIDAD 2: SERIES TEMPORALES

---

**Emilio Barragán Rodríguez**

Asignatura: 06MBID Estadística avanzada

Curso: 2023-24

Grupo: B

## Índice

- [1 Introducción](#)
  - [1.1 Contexto y motivación](#)
  - [1.2 Objetivos del análisis](#)
- [2 Descripción de los datos a analizar](#)
  - [2.1 Características estadísticas generales](#)
  - [2.2 Descomposición de la serie temporal](#)
    - [Tendencia](#)
    - [Estacionalidad](#)
    - [Variaciones aleatorias](#)
- [3 Análisis](#)
  - [3.1 Análisis de autocorrelación](#)
  - [3.2 Aplicación de un modelo adecuado](#)
- [4 Conclusiones](#)
  - [4.1 Resultados](#)
  - [4.2 Limitaciones](#)
- [5 Anexo](#)
  - [5.1 Código](#)

# 1 Introducción

En este apartado se explicará el contexto y la motivación de la actividad. También se describirán los objetivos de la misma.

## 1.1 Contexto y motivación

Las series temporales están en el día a día. Desde el precio del aceite hasta el número de accidentes de tráfico diarios. Es por ello que en esta actividad se trabajará haciendo uso de una serie temporal. En concreto, con datos del precio diario del oro.

## 1.2 Objetivos del análisis

El objetivo principal es aplicar lo aprendido en la asignatura de estadística. Estos objetivos incluirían:

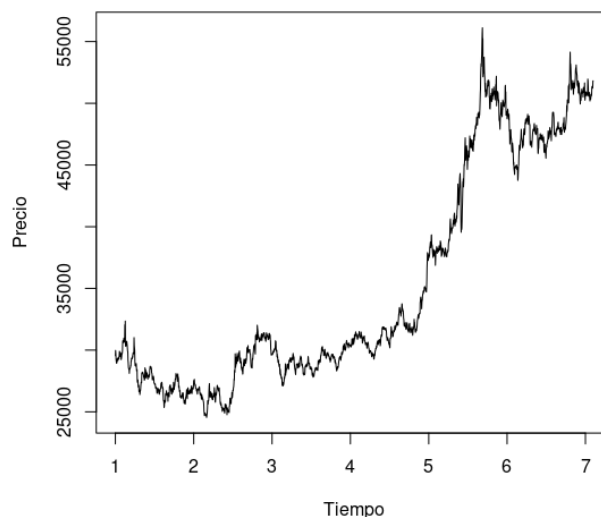
- Teoría sobre series temporales (descomposición, autocorrelación, modelos, etc).
- Programación en R.

# 2 Descripción de los datos a analizar

Los datos son sobre el precio diario del oro. Desde el día 1 de enero de 2014 hasta el día 5 de agosto de 2022. El dataset se puede encontrar en el siguiente enlace:

<https://github.com/emibarrod/estadistica-viu/blob/main/Actividad2/Gold%20Price.csv>

La serie temporal tiene el siguiente aspecto:



## 2.1 Características estadísticas generales

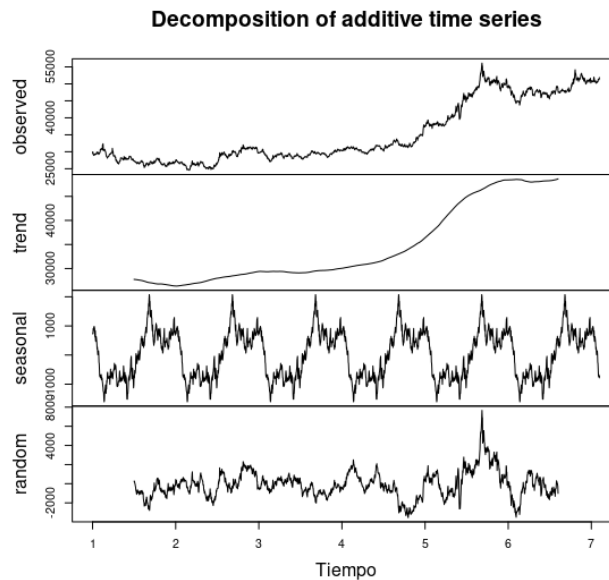
La frecuencia con la que se creará la serie temporal será anual (365 días). Esto es debido a la gran cantidad de datos que componen la serie y a que se analizará el precio del oro de forma anual.

- Número de registros: 2227
- Precio mínimo del oro: 24545

- Primer cuartil: 28554
- Mediana: 30766
- Precio medio del oro: 35218
- Tercer cuartil: 45942
- Precio máximo del oro: 56117

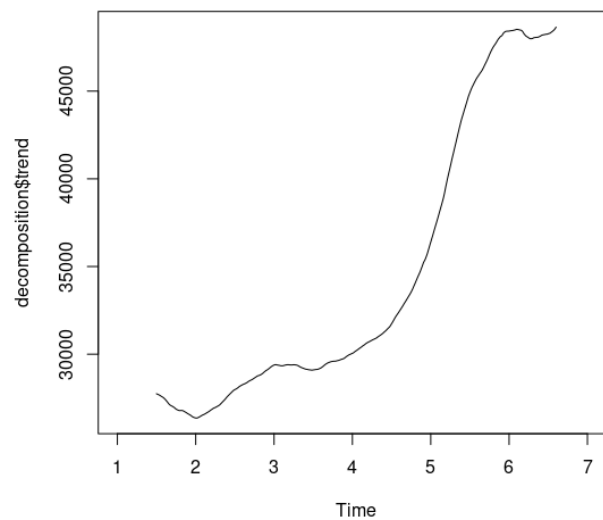
## 2.2 Descomposición de la serie temporal

Usando una descomposición aditiva se obtienen los siguientes componenetes de la serie:



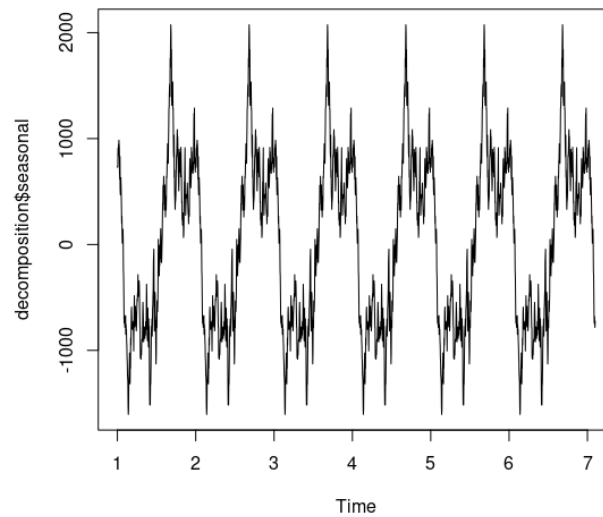
### Tendencia

Se puede observar que la tendencia es positiva y bastante marcada.



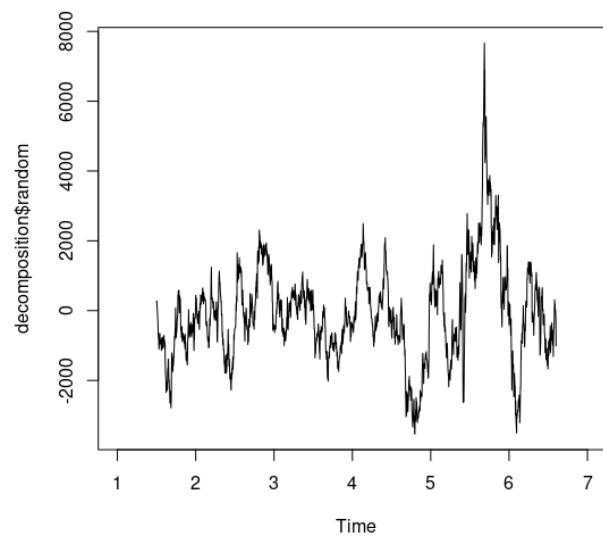
### Estacionalidad

En la gráfica de la estacionalidad se aprecia un patrón anual.



## Variaciones aleatorias

Las variaciones aleatorias han sido constantes en la serie, teniendo un pico en el sexto año de la serie.

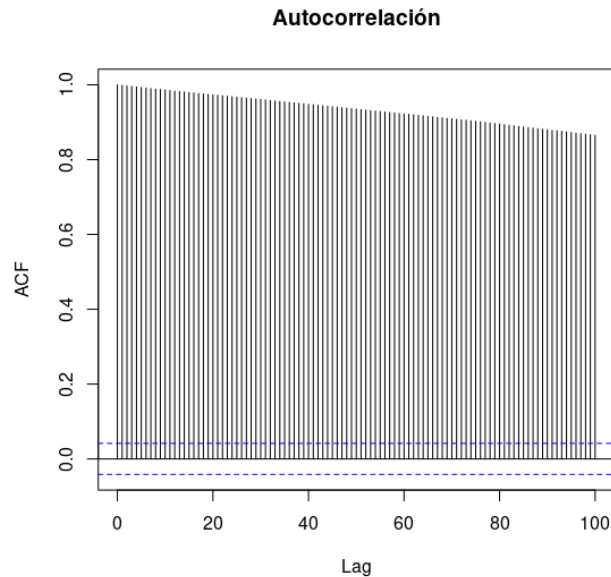


## 3 Análisis

En este apartado se analizará la autocorrelación de la serie. También se construirá un modelo predictivo para la serie.

### 3.1 Análisis de autocorrelación

Al calcular la autocorrelación de la serie para un lag de hasta 100 unidades, se obtiene:



Esto indica que la serie depende mucho de sí misma, pero cuanto más lejos en tiempo el dato, menos relevante es.

### 3.2 Aplicación de un modelo adecuado

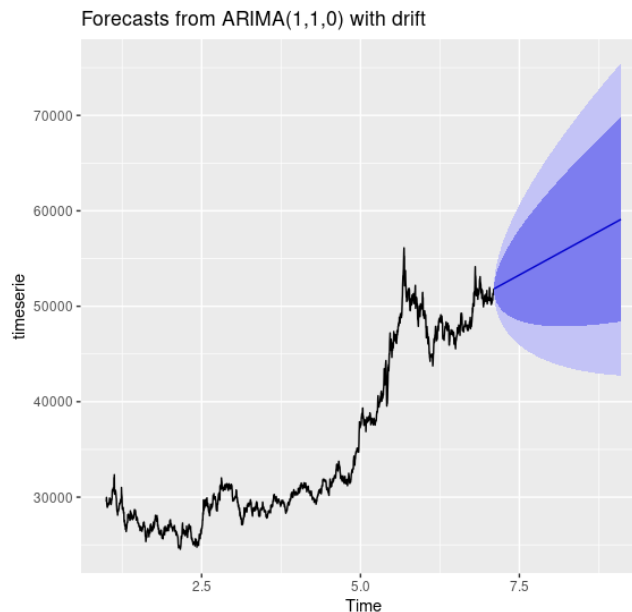
Se ha elegido aplicar el modelo ARIMA. Para obtener el modelo ARIMA óptimo se ha usado el método auto ARIMA. El modelo conseguido tiene la siguiente parametrización:

```
ARIMA(1,1,0) with drift

Coefficients:
      ar1    drift
    -0.0424  9.9834
s.e.    0.0212  6.5483

sigma^2 = 103811:  log likelihood = -16013.07
AIC=32032.14    AICc=32032.15    BIC=32049.27
```

Haciendo uso de este modelo, se obtienen las siguientes predicciones:



Como se puede apreciar, el modelo predice una subida del precio del oro.

## 4 Conclusiones

En este apartado se presentarán los resultados y las limitaciones de nuestro estudio.

### 4.1 Resultados

Como se puede apreciar, la serie temporal se comporta de manera similar a lo largo del tiempo. Tendencia positiva y sin muchas oscilaciones.

Por otra parte, se ve que nuestro modelo consigue imitar ese comportamiento para las predicciones.

### 4.2 Limitaciones

La primera limitación son los datos. Aunque puedan parecer muchos datos, si queremos realizar un análisis exhaustivo del comportamiento anual del precio del oro, habría que conseguir más datos.

Por otra parte, se podrían probar otros modelos estadísticos, como el SARIMA. Esto nos lleva también al punto de que se podría eliminar la tendencia y la estacionalidad de la serie, así como las variaciones aleatorias.

En resumen, se ha conseguido un estudio básico de una serie temporal. También se ha creado un modelo adecuado para la serie. No obstante, se podrían aplicar más técnicas y métodos para el estudio.

## 5. Anexo

### 5.1 Código

```
# instalación y carga de las librerías necesarias

install.packages('anytime')
install.packages('forecast', dependencies = TRUE)
library(anytime)
```

```

library(forecast)

# Preparación del dataset

gold_data <- read.csv(
  "/home/emibarrodd/Documentos/Actividad2_estadistica/Gold Price.csv"
)
gold_data <- subset(gold_data, select = c("Date", "Price"))
View(gold_data)

# Creación de la serie temporal

timeserie <- ts(gold_data$Price, frequency = 365)
timeserie
plot.ts(timeserie, xlab="Tiempo", ylab = "Precio")

# Estadísticas de la serie temporal

length(timeserie)
start(timeserie)
end(timeserie)
head(timeserie)
tail(timeserie)
summary(timeserie)

# Descomposición

decomposition <- decompose(timeserie)
plot(decomposition, xlab="Tiempo")
plot(decomposition$trend)
plot(decomposition$seasonal)
plot(decomposition$random)

# Autocorrelación

autocorr <- acf(as.numeric(timeserie), 100, plot=FALSE)
plot(autocorr, main="Autocorrelación")
autocorr

# Auto ARIMA

fit_auto_arima <- auto.arima(timeserie)
fit_auto_arima
forecast(fit_auto_arima)
autoplot(forecast(fit_auto_arima))

```