

# Hoja informativa: Análisis de series temporales

## Práctica

```
# reconocimiento de fechas y formación de nuevos índices
# valores de index_col = la lista de números de columna o nombres de columna
# valores de parse_dates = la lista de números de columna o nombres de columna
data = pd.read_csv('filename.csv', index_col=[0], parse_dates=[0])
```

```
# comprobar que el índice es monótono
print(data.index.is_monotonic)
```

```
# remuestreo: media para cada hora
data.resample('1H').mean()

# remuestreo: máximo por cada dos semanas
data.resample('2W').max()
```

```
# media móvil con tamaño de ventana = 7
data.rolling(7).mean()
```

```
# descomposición de la serie temporal en tendencia, estacionalidad y residuos
from statsmodels.tsa.seasonal import seasonal_decompose

decomposed = seasonal_decompose(data)

decomposed.trend # tendencia
decomposed.seasonal # estacionalidad
decomposed.resid # residuos
```

```
# cambio de un paso con el llenado del valor cero
print(data.shift(fill_value=0))
```

## Teoría

Las **series temporales** son las secuencias de números a lo largo del eje del tiempo. El intervalo entre los valores de la serie es constante.

**Remuestrear** significa cambiar el intervalo con los valores de la serie. Se realiza en dos pasos:

1. Elige la nueva duración del intervalo. Considera que los valores del intervalo existente están agrupados.
2. En cada grupo se calcula el valor acumulado de la serie. Puede ser mediana, media, máximo o mínimo.

La **media móvil** o **promedio móvil** es un método para suavizar los datos en una serie temporal. El método consiste en encontrar los valores menos susceptibles a fluctuaciones, es decir, la media aritmética.

Una **tendencia** es un cambio ligero del valor medio de la serie sin repetir patrones.

**Estacionalidad** significa patrones que se repiten de forma cíclica en una serie temporal.

El **proceso estocástico** tiene una variación aleatoria y su distribución cambia con el tiempo.

Un proceso estocástico es **estacionario** si su distribución no cambia con el tiempo. Si la distribución cambia, entonces el proceso *estocástico* es **no estacionario**.

Las **diferencias de series temporales** son una secuencia de diferencias entre elementos vecinos de una serie temporal (es decir, el valor anterior se resta del siguiente).