

# Actividad 5

Johana Guadalupe Miranda González

Febrero 2021

## Introducción

El objetivo de esta actividad es usar los mismos datos climatológicos que hemos estado utilizando hasta el momento, explorar un análisis de datos en series de tiempo para observar de mejor manera los datos.

En esta actividad repasamos el análisis de datos de la paquetería **Pandas**. De igual manera, agregaremos dos paqueterías extra: **Seaborn** y **statmodels.api**.

## Descripción

Lo primero que haremos es crear un nuevo cuaderno trabajo de Jupyter y en la primera celda vamos a concentrar todas las bibliotecas que utilizaremos para que no se pierdan más abajo. Después en la segunda celda, realiza todas las operaciones sobre el DataFrame de datos original hasta terminar con el DataFrame `df_meteorologicos` que tiene las variables como se desean.

Después, creamos un conjunto de datos de los últimos años a partir de 1990 y exploramos los datos de precipitación, temperatura máxima, temperatura mínima y evaporación.

Utilizaremos la función **`df.plot()`** para graficar el rango de datos utilizados, la función **`df.loc()`** para analizar un año específico completo y la función **`sns.barplot()`** para graficar la estacionalidad de tus datos por mes.

Una vez que analizamos lo anterior, creamos un nuevo DataFrame de las Temperaturas y la Evaporación con promedios móviles de 7, 30 y 365 días, utilizando la función **`df.rolling()`**. Hecho esto, analizaremos un año de datos, contrastando los datos diarios con el promedio móvil de 7 y 30 días. Después exploraremos para 10 o más años de datos, contrastando los datos diarios con las gráficas de promedio móvil de 365 días.

También utilizaremos la biblioteca **statsmodels.tsa** de Python, para realizar una descomposición de una serie de tiempo: Serie Observada = Tendencia + Estacionalidad + Residuo.

Ya por último, exploraremos visualmente si se observa un cambio en la distribución de Temperaturas (histogramas y kde), sobreponiendo dos histogramas, en mi caso de dos periodos de 20 años contiguos, debido a que mis datos empiezan en 1975.

## Conclusión

En esta actividad, creo que me familiarice mucho más con la creación de DataFrame y con las gráficas. Considero que estas herramientas son muy útiles para hacer análisis de datos. Me gusto esta actividad, se me hizo un poco larga pero entretenida.