



Universidad de Sonora  
División de Ciencias Exactas y Naturales  
Departamento de Física  
Física computacional  
**Análisis de Series de Tiempo con Pandas**  
**Actividad 5**  
Iveth R. Navarro

12 de febrero de 2021  
Puerto Peñasco, Sonora, México

### Resumen

La presente es una práctica de Física Computacional cuyo objetivo es el empleo de Análisis de Series de Tiempo en Python siendo la información meteorológica de CONAGUA la estudiada, estación Puerto Peñasco, Sonora.

## 1. Introducción

Una serie de tiempo es cualquier conjunto de datos ordenados en el tiempo a intervalos iguales. Hemos visto que los datos climatológicos que descargamos del Servicio Meteorológico Nacional son datos diarios, se tiene una serie de tiempo para la Precipitación, una para la Evaporación, y las de las Temperaturas (Máxima y Mínima). Podrían ser datos cada hora, cada minuto o quizás 10 datos por segundo, dependiendo del fenómeno que se deseé estudiar.

Se llama Análisis de Series de Tiempo al proceso de explorar patrones, regularidades, periodicidad, o realizar un análisis estadístico de una serie de tiempo (promedio aritmético, desviación estándar, promedios móviles, otras).

Hay varios técnicas para analizar Series de Tiempo:

- Una es en el dominio temporal (auto-correlación, correlación-cruzada, tendencia, estacionalidad, descomposición de series, otras).
- Otra es en el espacio de frecuencias (análisis espectral, análisis de ondículas (wavelets)). Por ejemplo: Análisis de Series de Fourier y Transformadas de Fourier.

En Física se utilizan mucho estas técnicas para el estudio de una gran diversidad de fenómenos, donde las variables física dependen del tiempo. Es un campo muy interesante, muy amplio y muy activo.

En nuestro caso nos limitaremos a hacer análisis en el dominio temporal ,y nos apoyaremos de Pandas para analizar nuestras series de tiempo de datos climatológicos.

## 2. Desarrollo

Esta actividad se dividió en 6 partes:

- Actividad 5.1: Explorar los datos de Precipitación los últimos años: 1990- fin de datos.
- Actividad 5.2: Explorar los datos de Temperaturas Máxima y Mínima de los últimos 30 años datos a partir de 1990.
- Actividad 5.3: Realizar el mismo análisis que en la Actividad 5.2, pero para la variable de Evaporación.
- Actividad 5.4: Crear nuevos DataFrames de las Temperaturas y la Evaporación con promedios móviles de 7, 30 y 365 días.

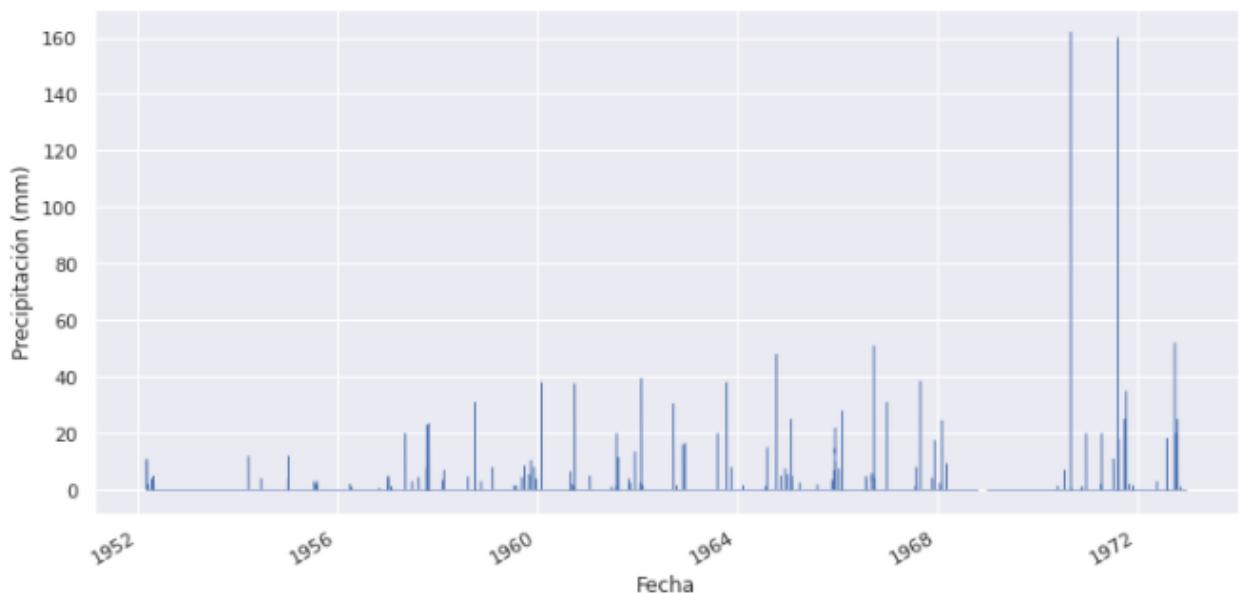
- Actividad 5.5: Utilizar la biblioteca *statsmodels.tsa* de Python, para realizar una descomposición de una serie de tiempo: *Serie Observada = Tendencia + Estacionalidad + Residuo*.
- Actividad 5.6: Explorar visualmente si se observa un cambio en la distribución de Temperaturas (histogramas y kde), sobreponiendo dos histogramas de dos períodos de 30 años contiguos.

Nota: En este caso, dada la cantidad de datos de Puerto Peñasco, los intervalos fueron de 20 años.

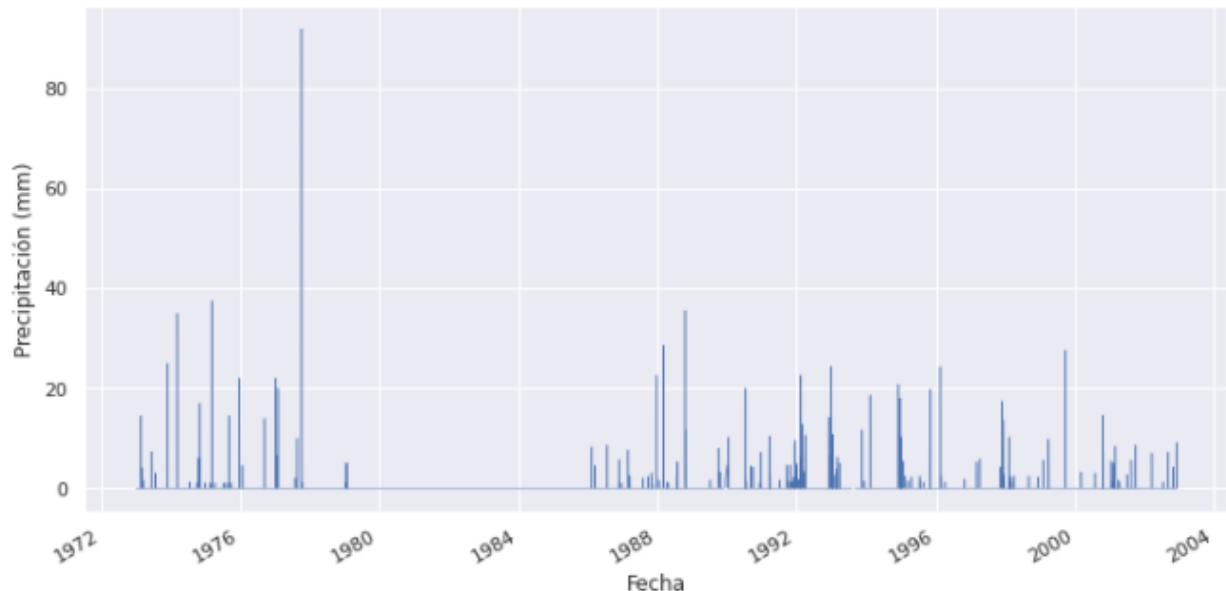
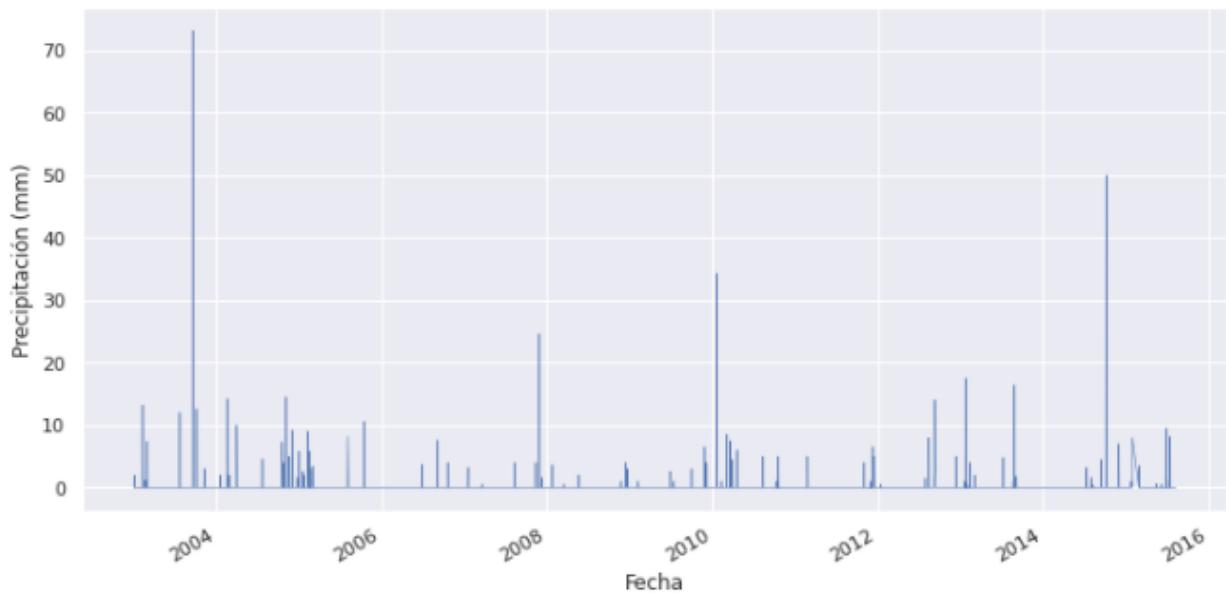
## 2.1. Resultados

### 2.1.1. Actividad 5.1

**Gráficas de todo el rango de datos.**



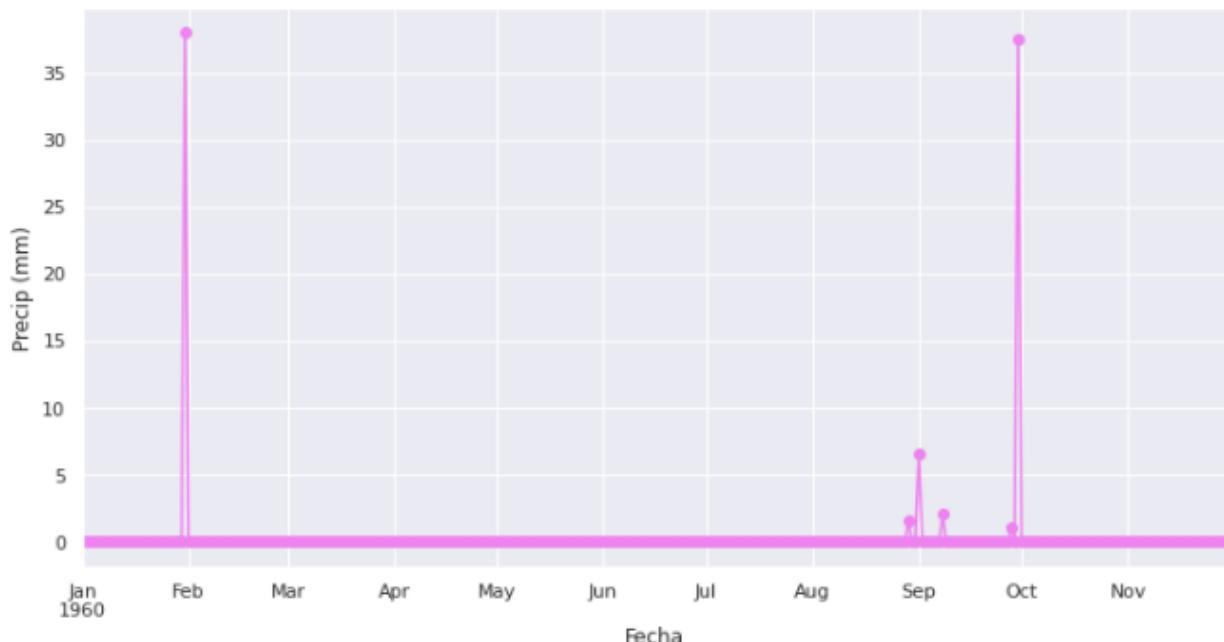
**Figura 1**

**Figura 2**

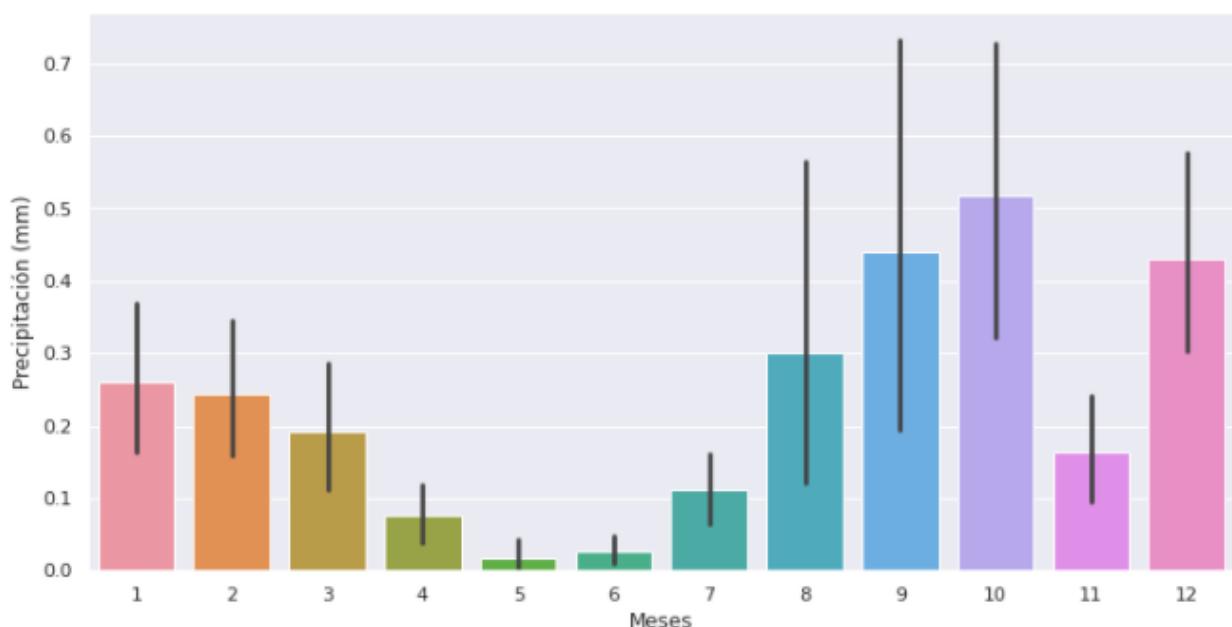
Variación en precipitación promedio 60-30: -0.10739177226651342  
Variación en precipitación promedio 90-60: -0.02270303154881917

**Figura 3**

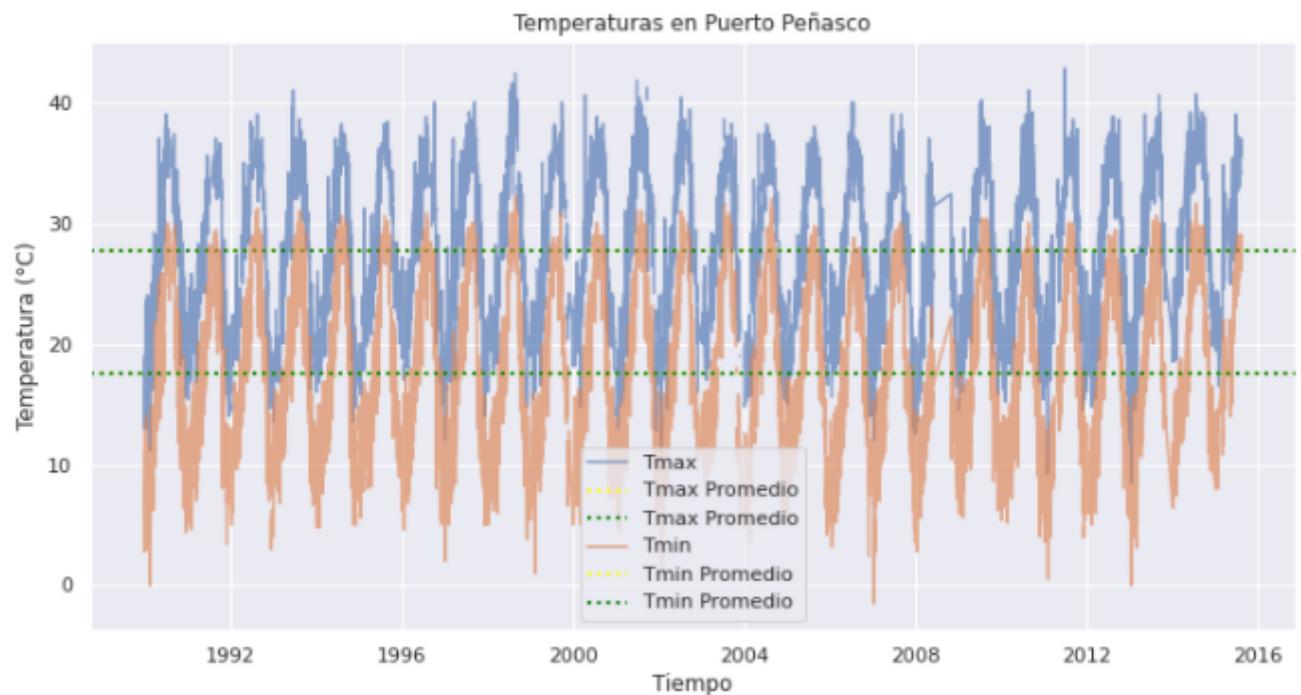
### Análisis del año 1960.

**Figura 4**

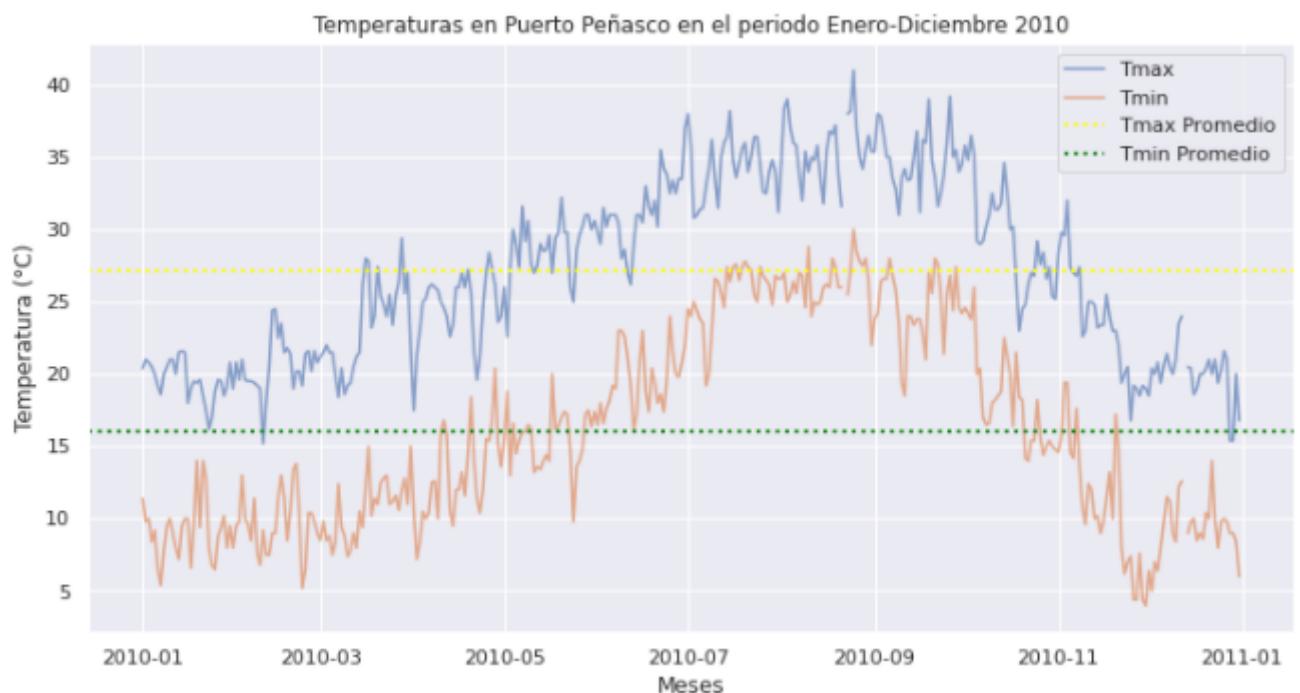
**Gráfica de la estacionalidad de los datos por mes.**

**Figura 5**

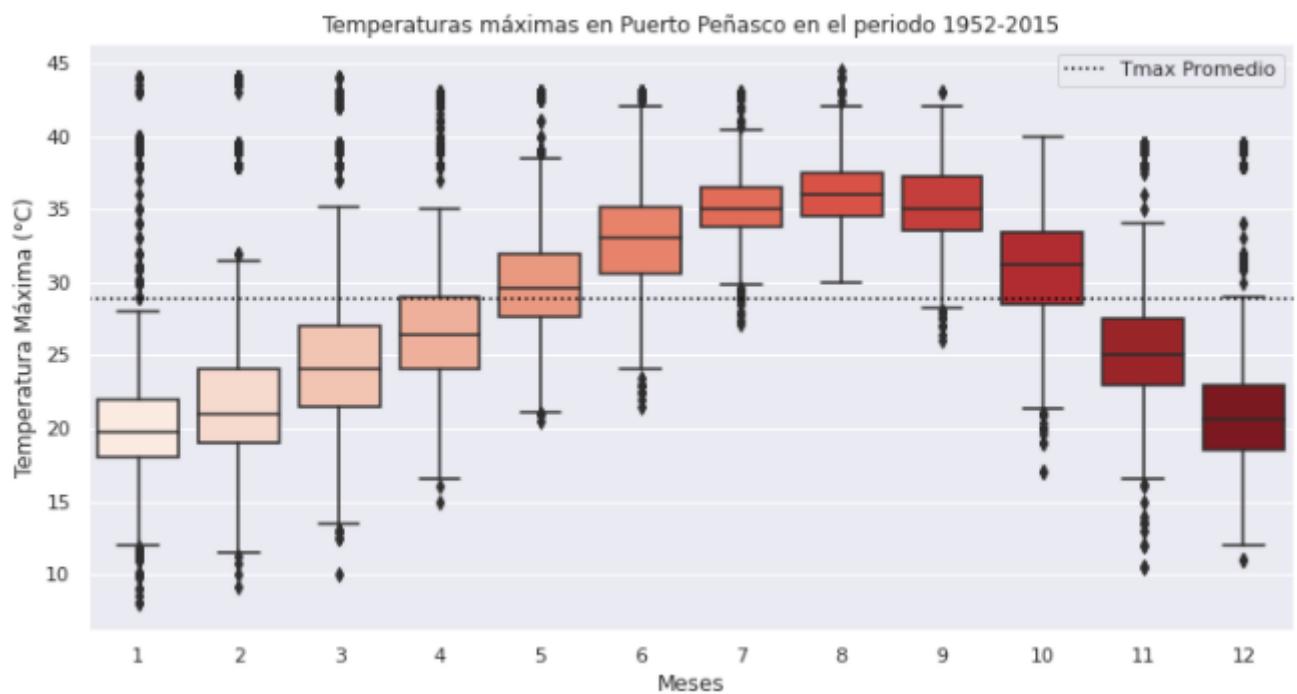
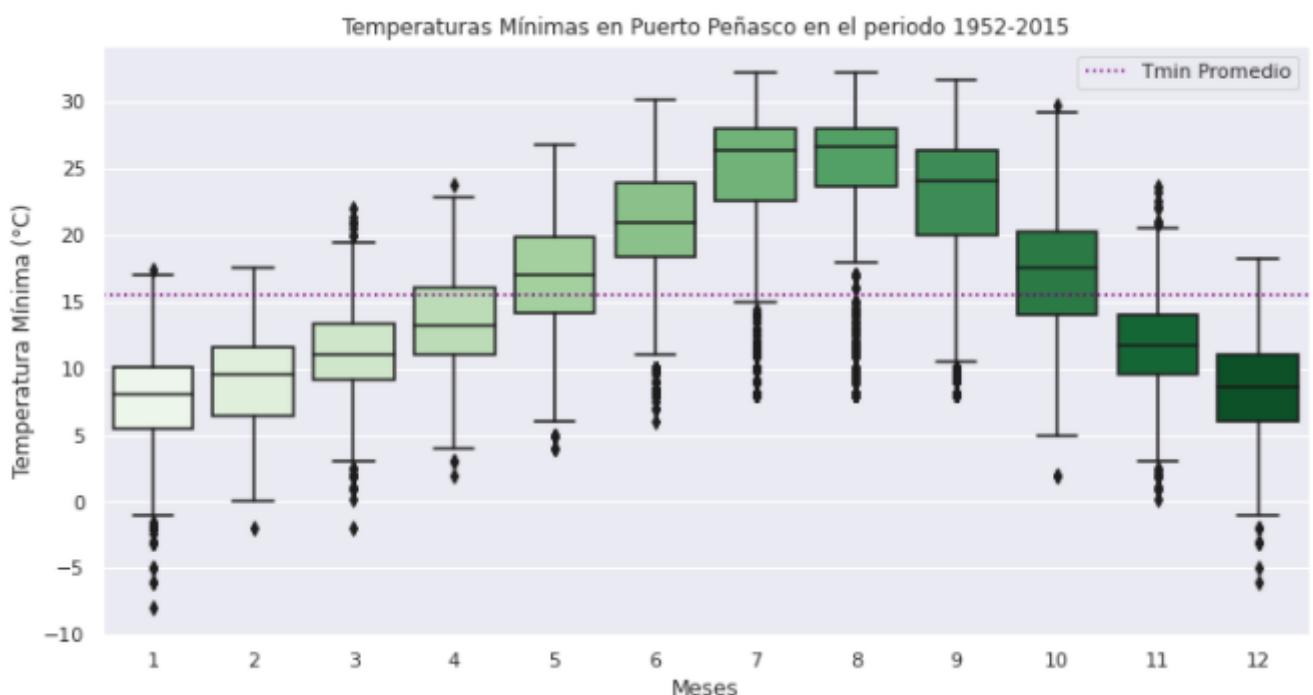
### 2.1.2. Actividad 5.2



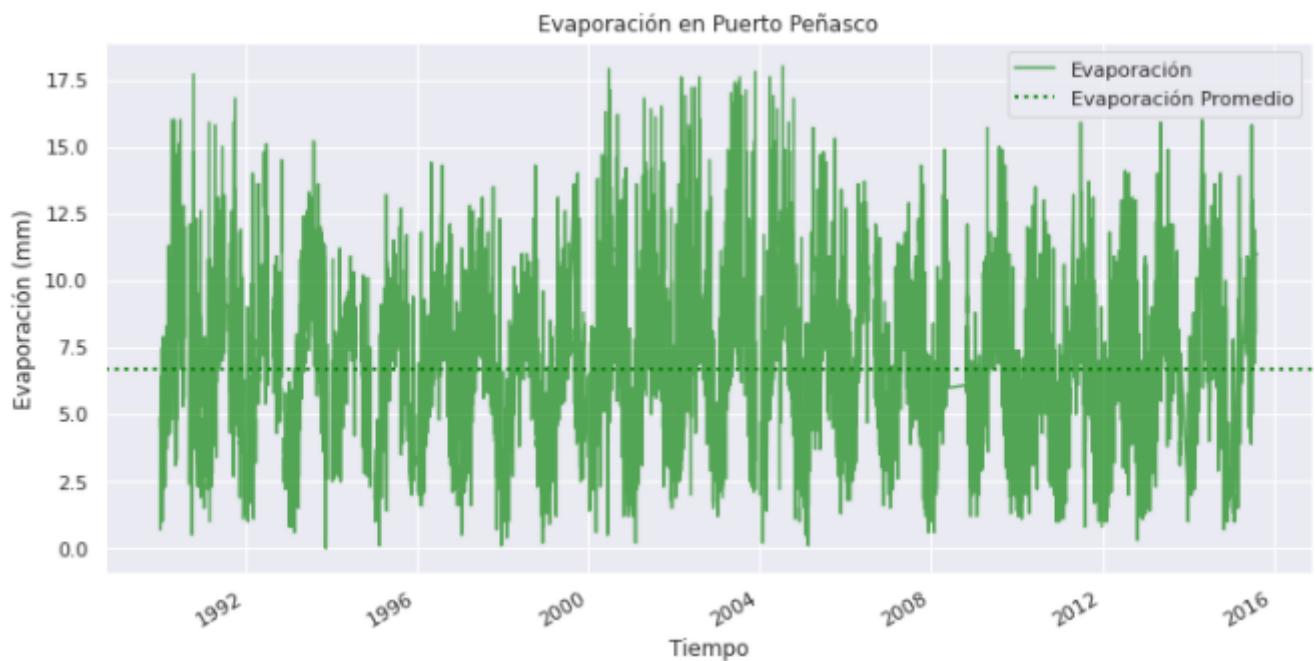
**Figura 6**



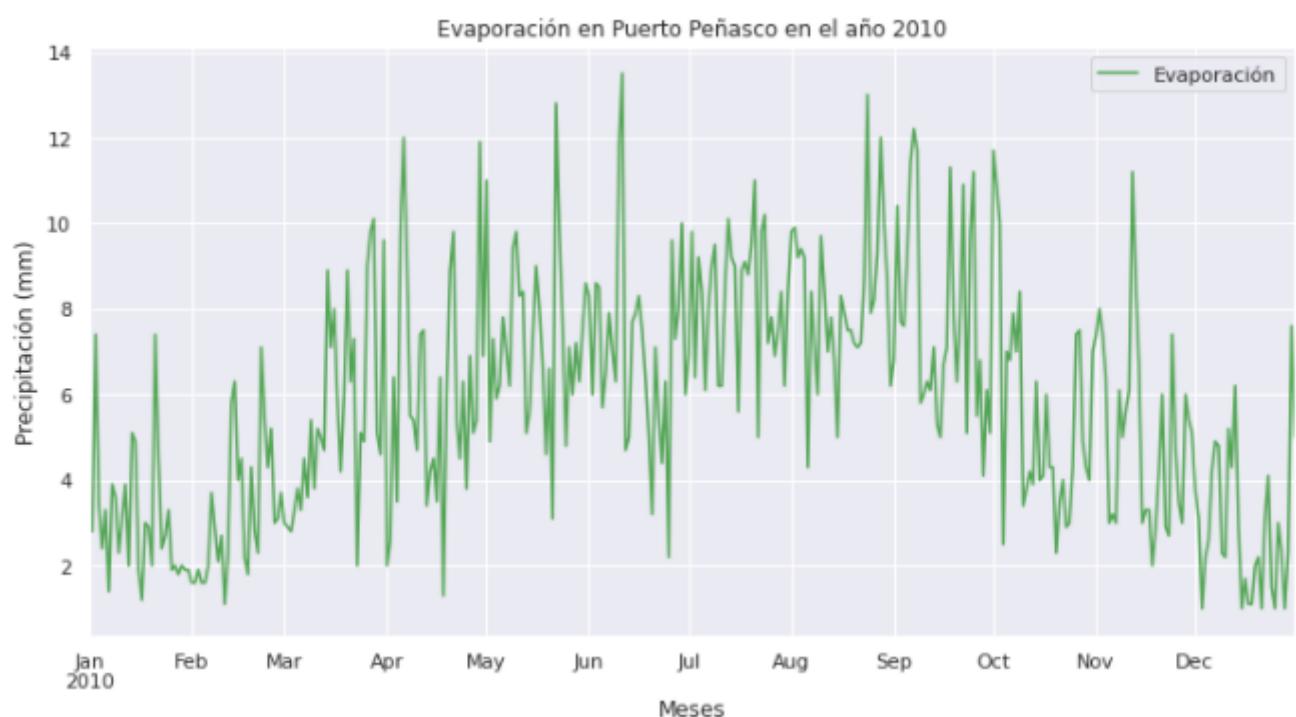
**Figura 7**

**Figura 8****Figura 9**

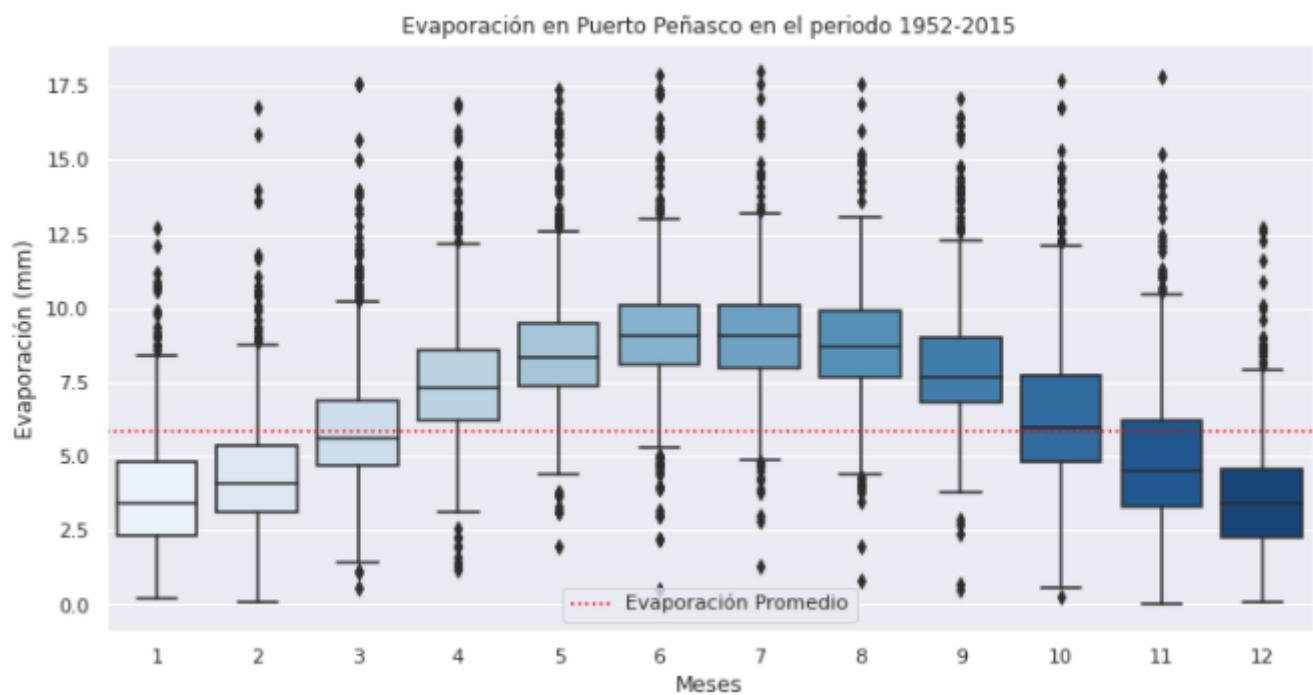
### 2.1.3. Actividad 5.3



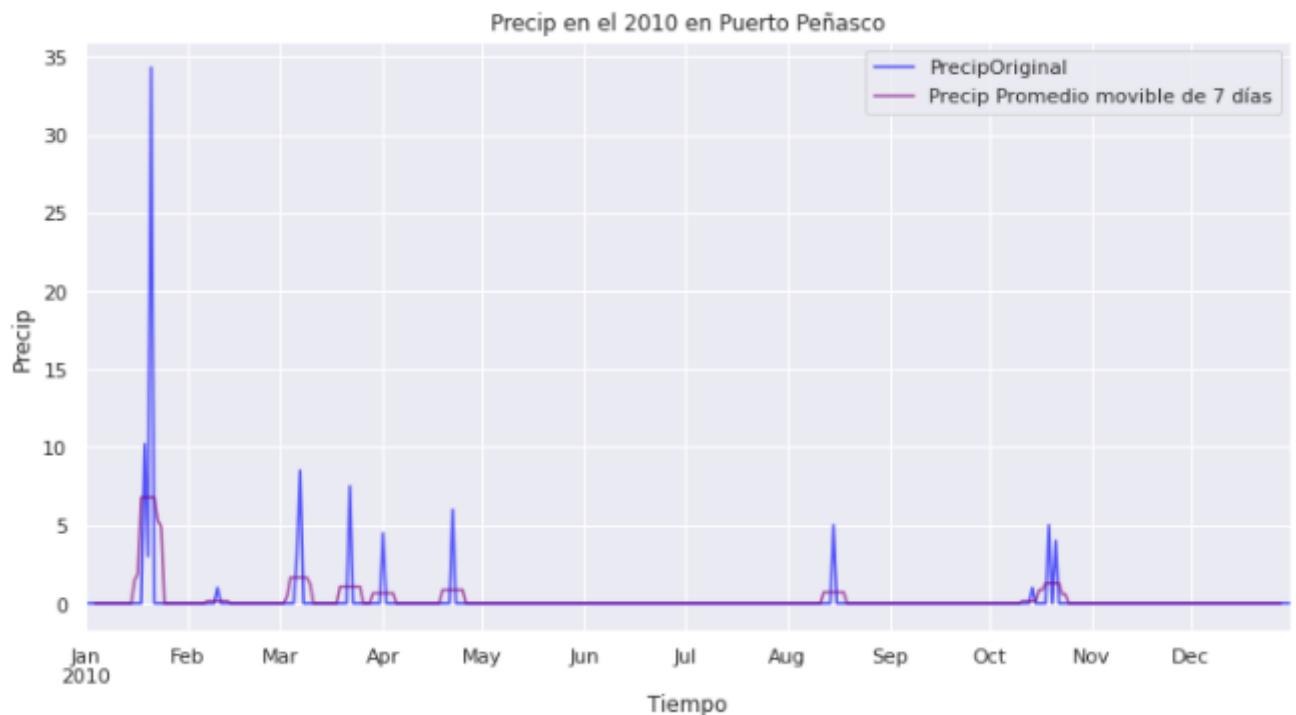
**Figura 10**

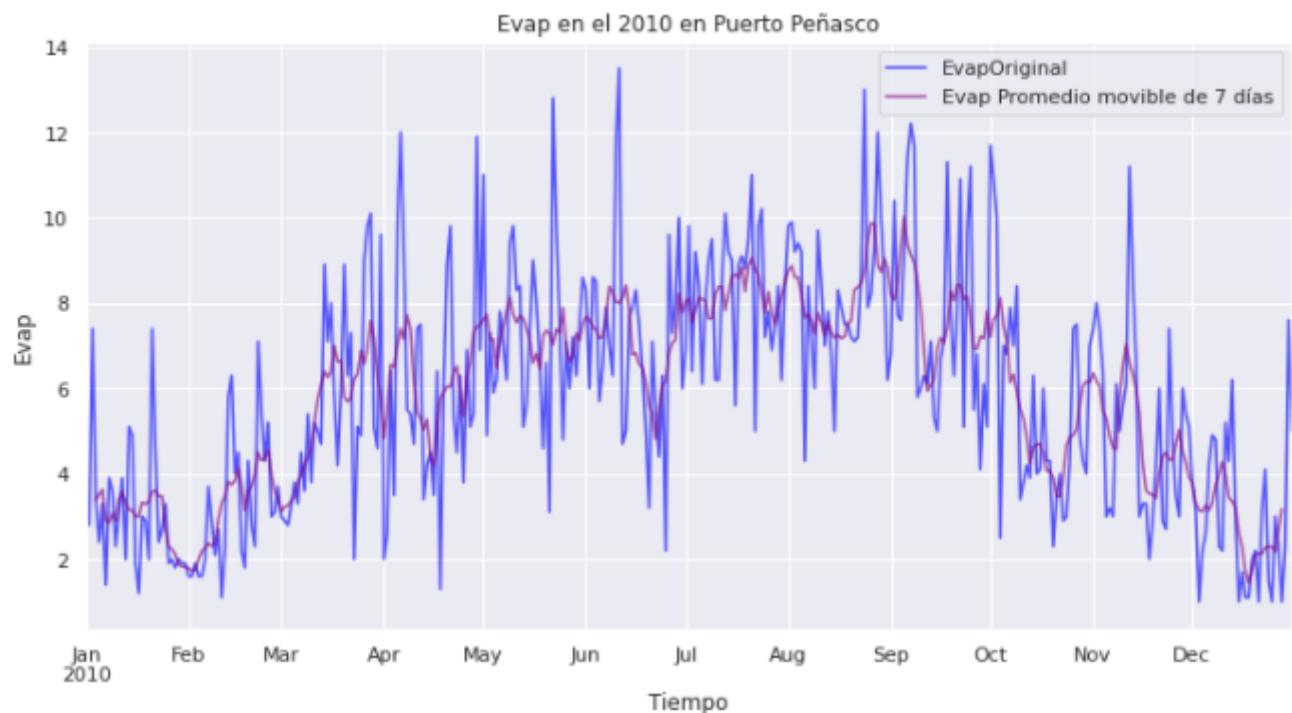


**Figura 11**

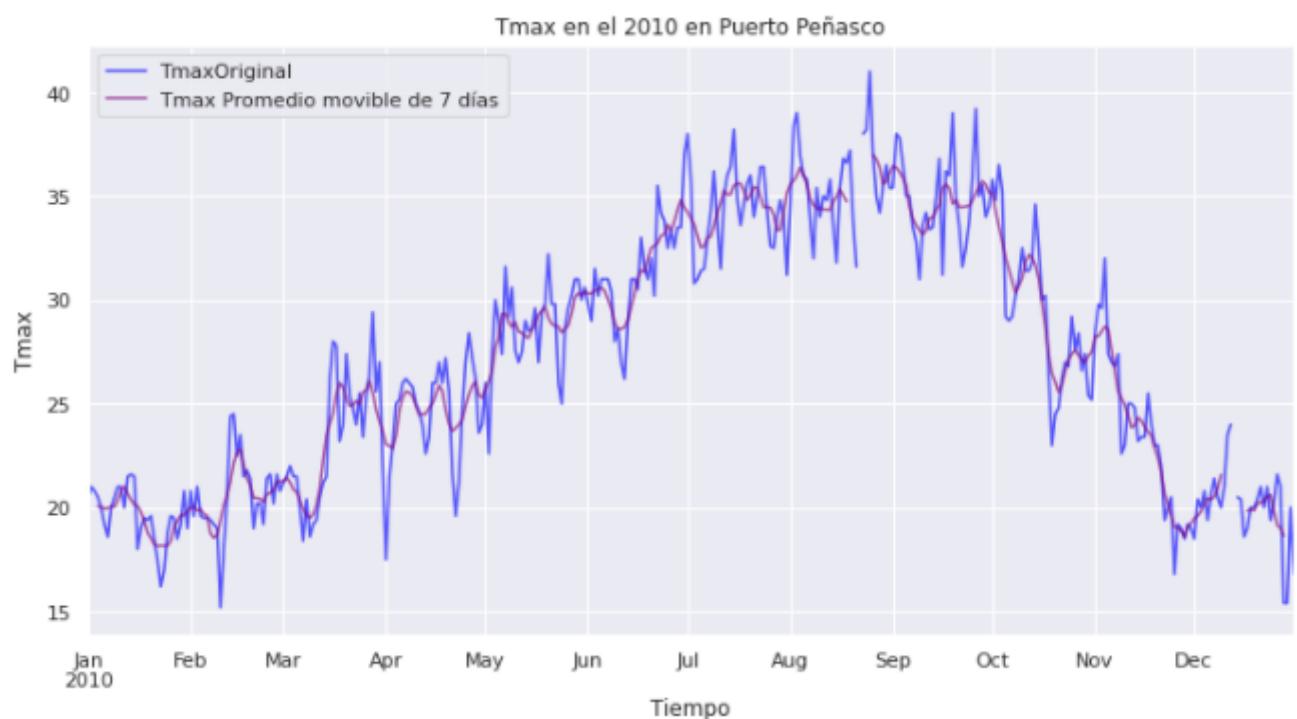
**Figura 12**

#### 2.1.4. Actividad 5.4

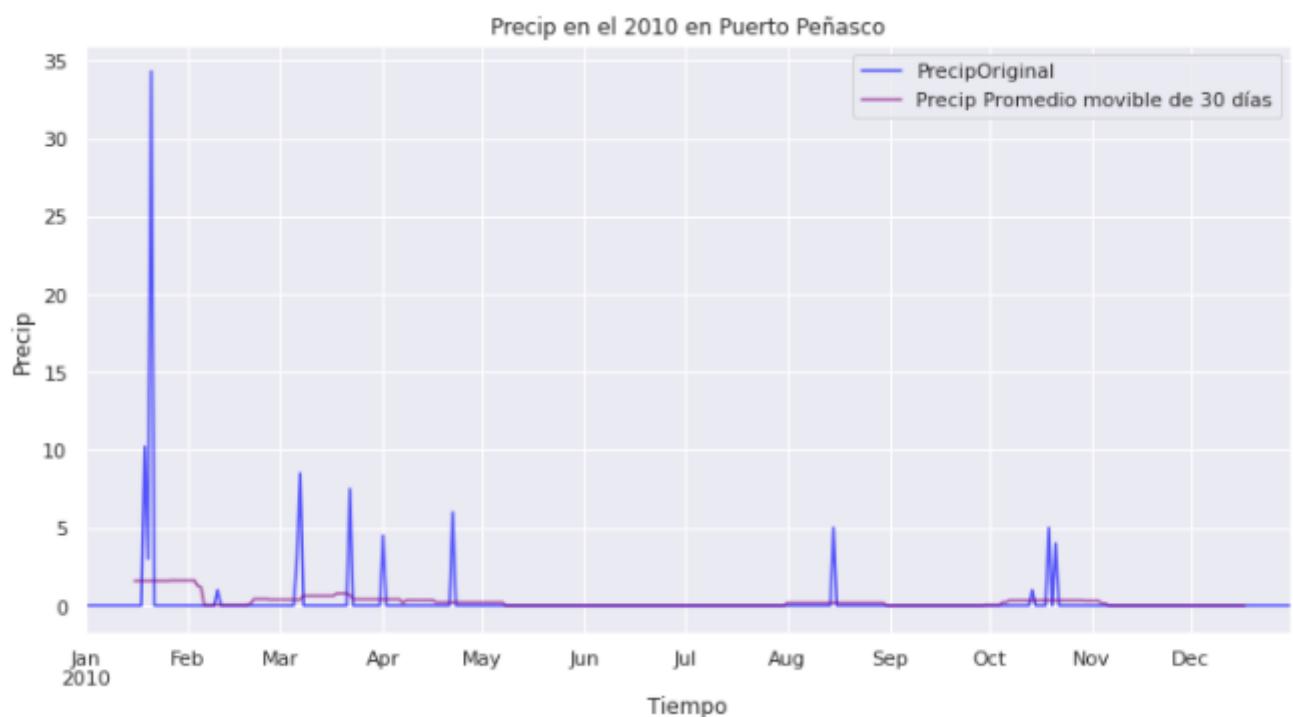
**Figura 13**

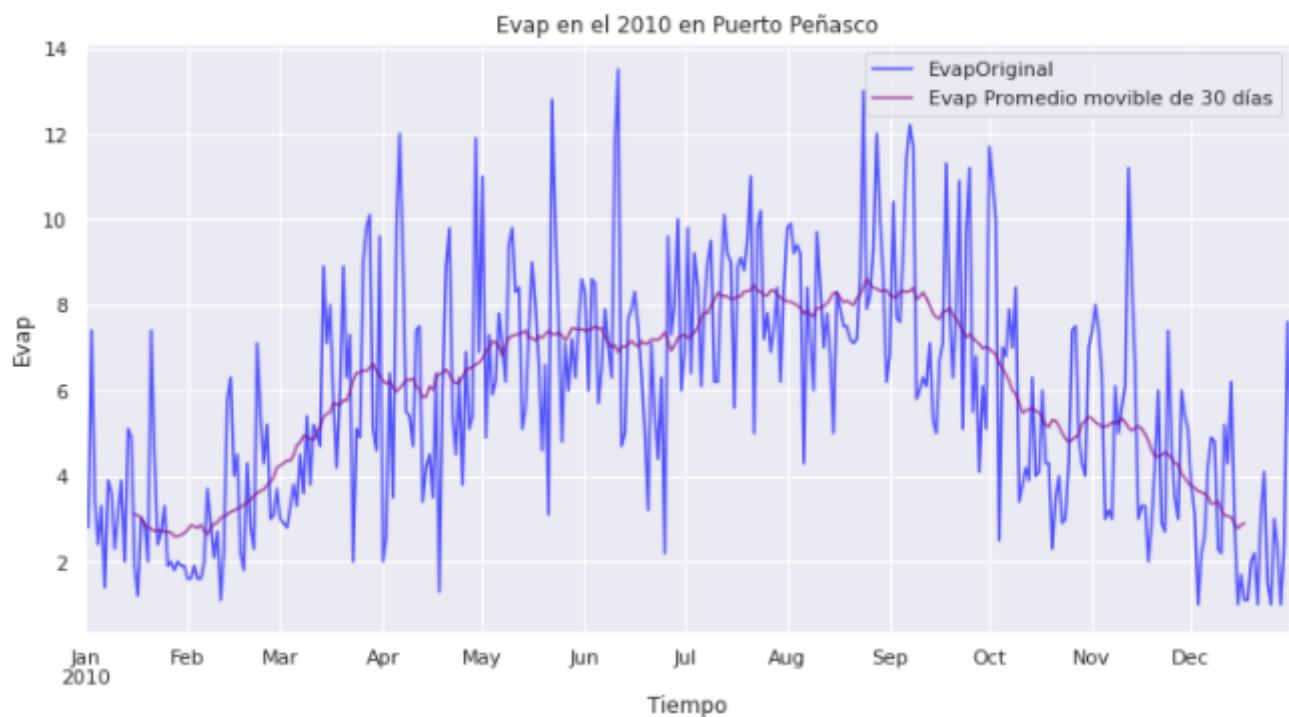
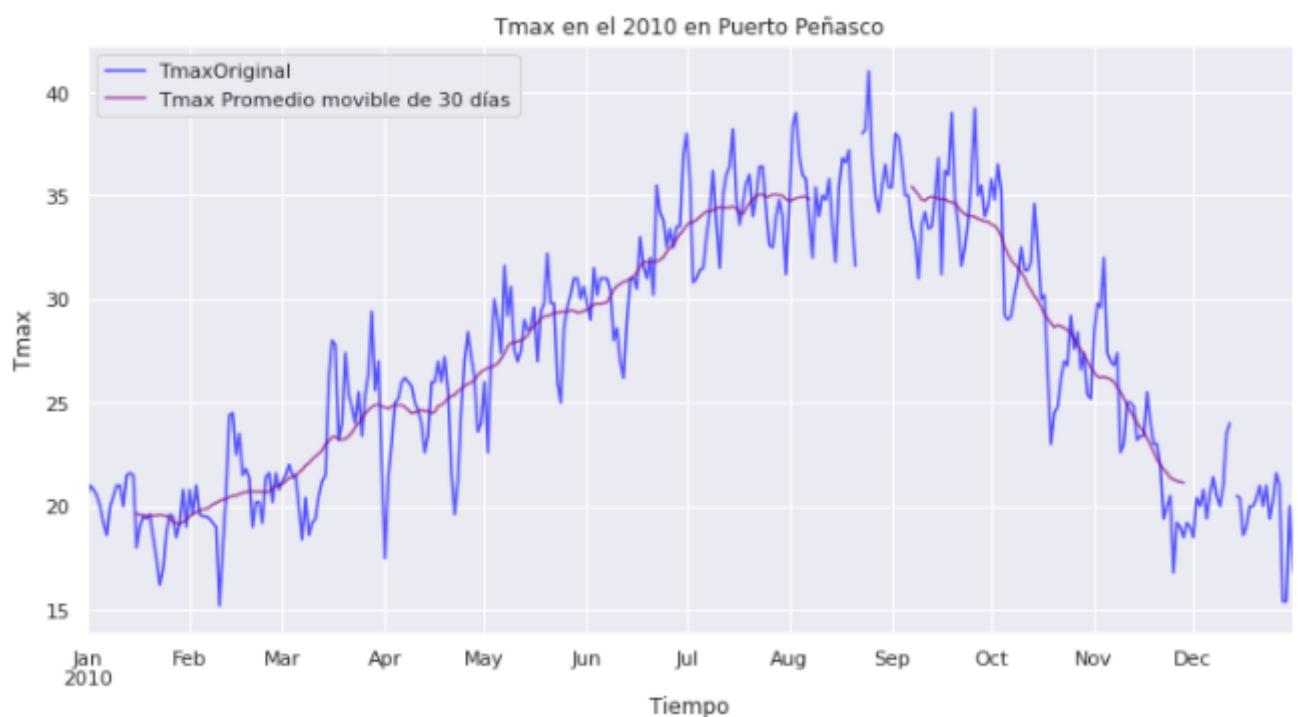


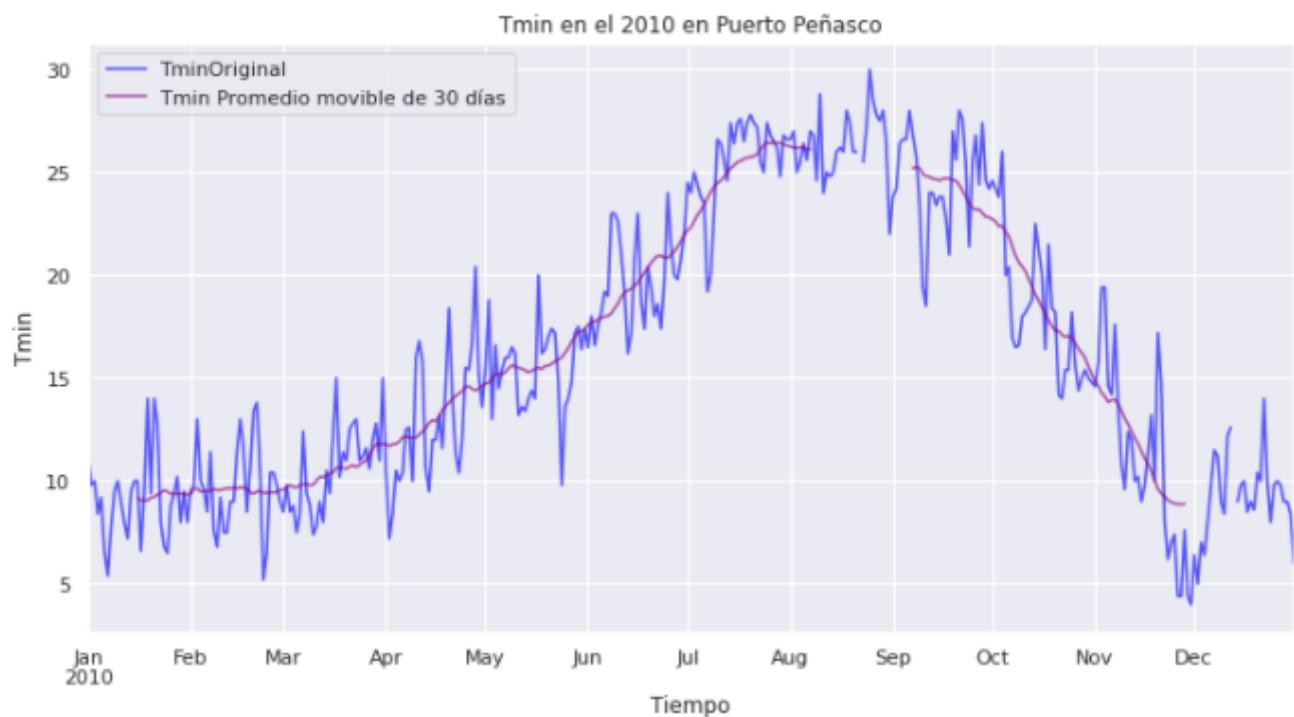
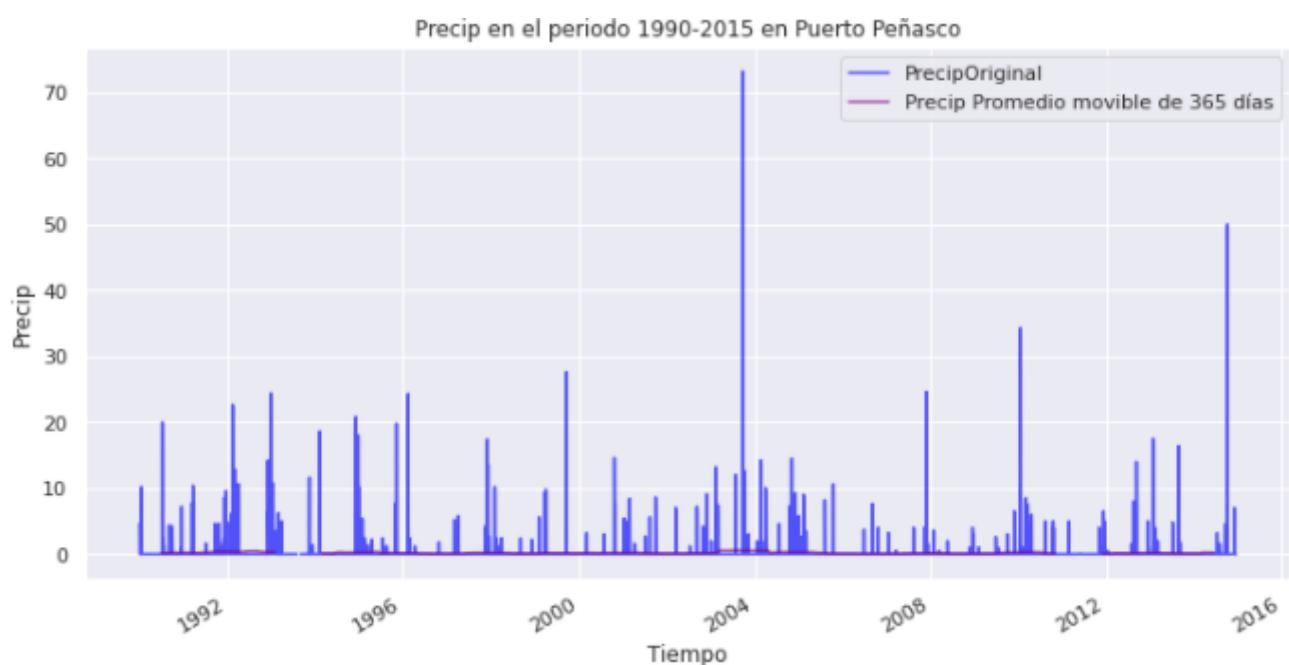
**Figura 14**

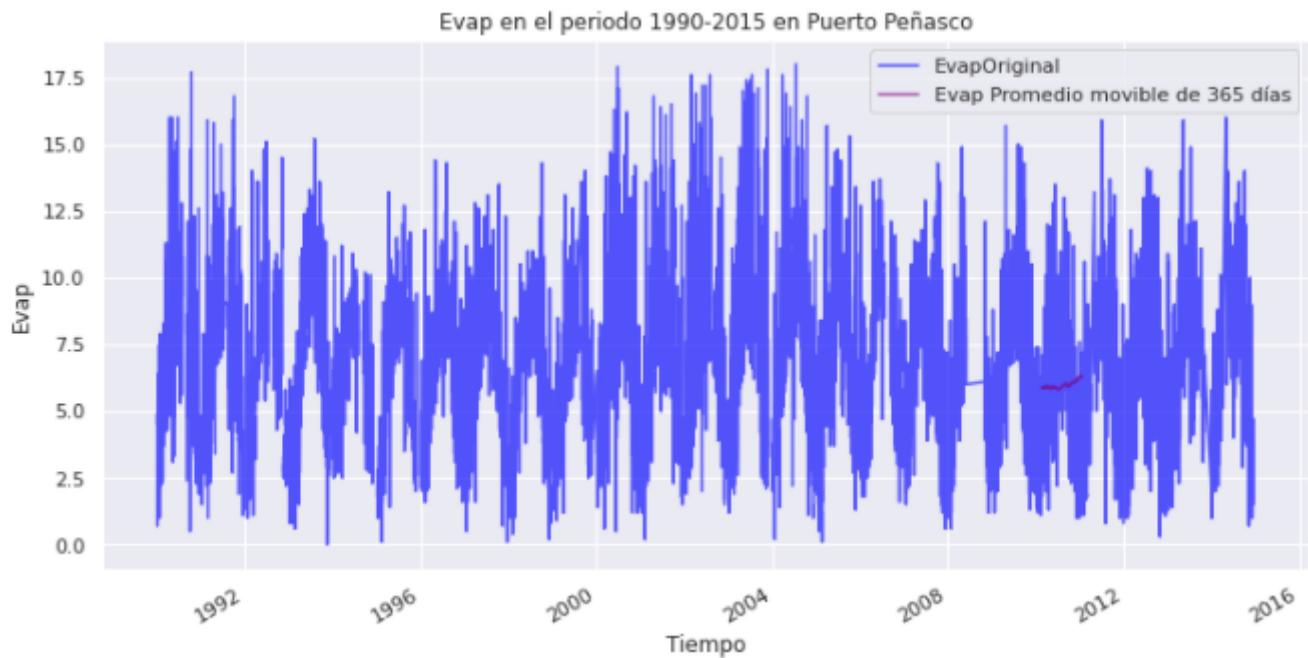
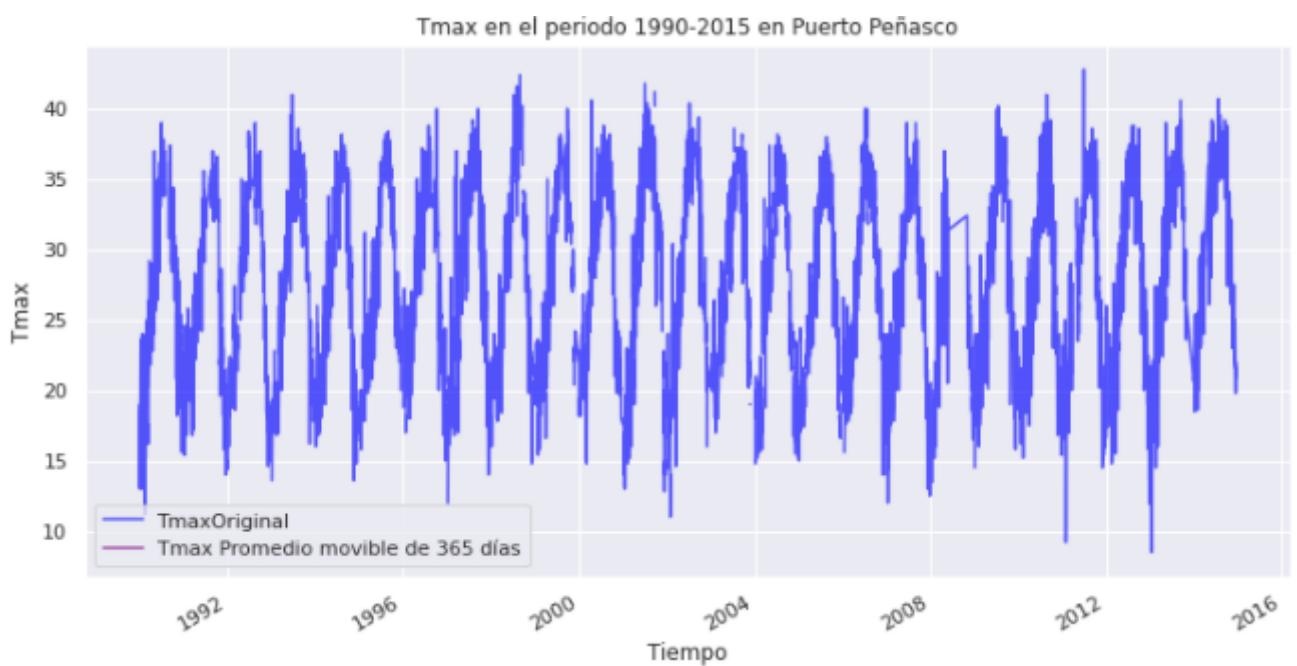


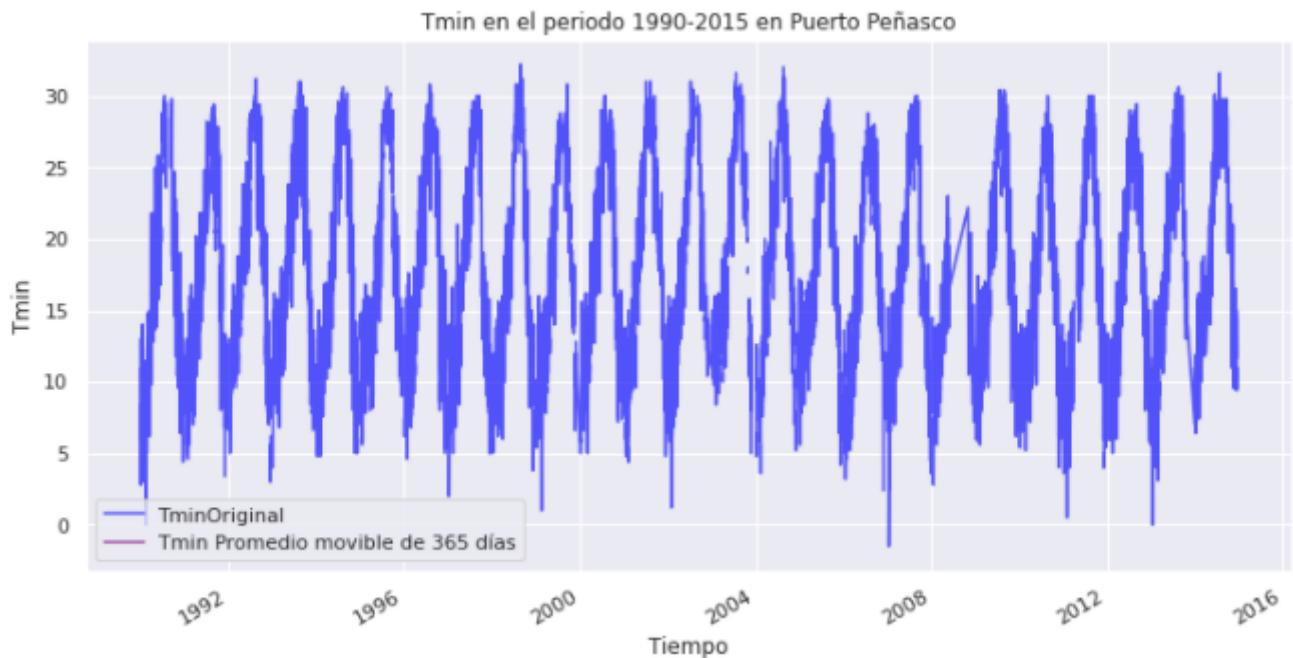
**Figura 15**

**Figura 16****Figura 17**

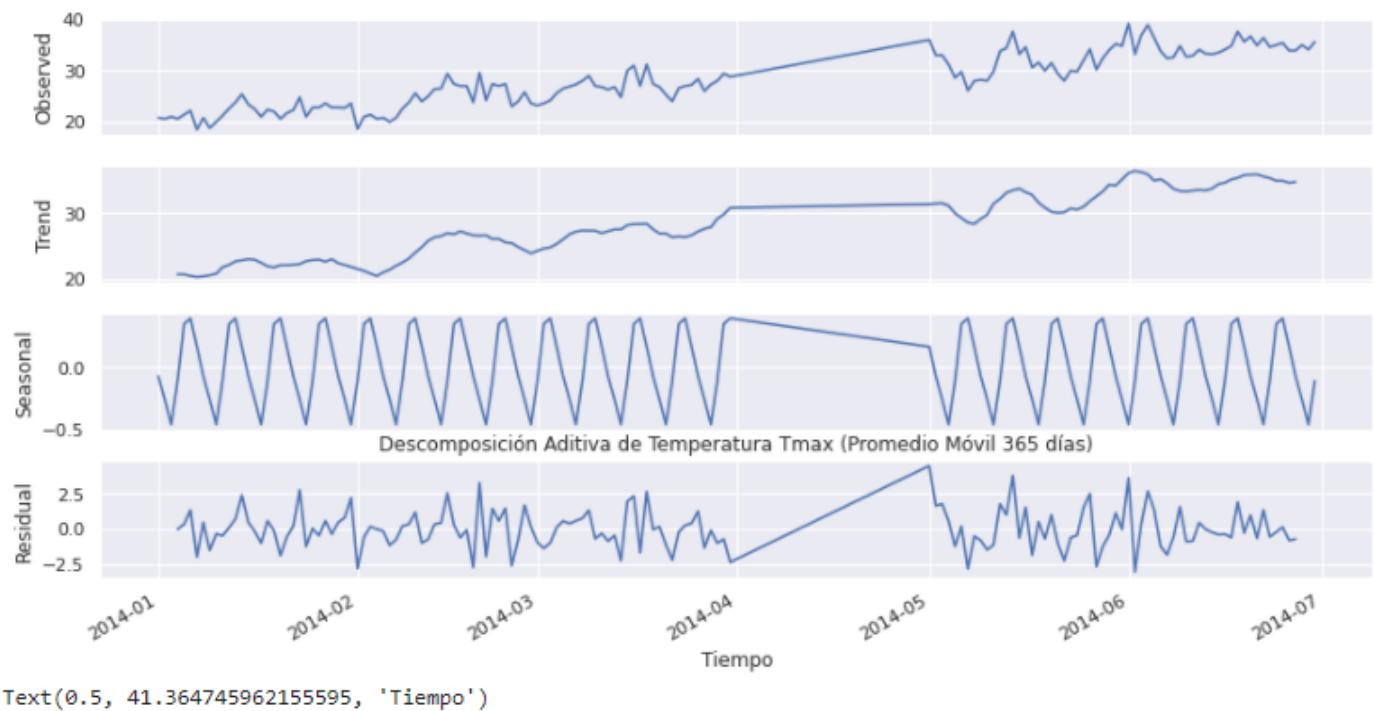
**Figura 18****Figura 19**

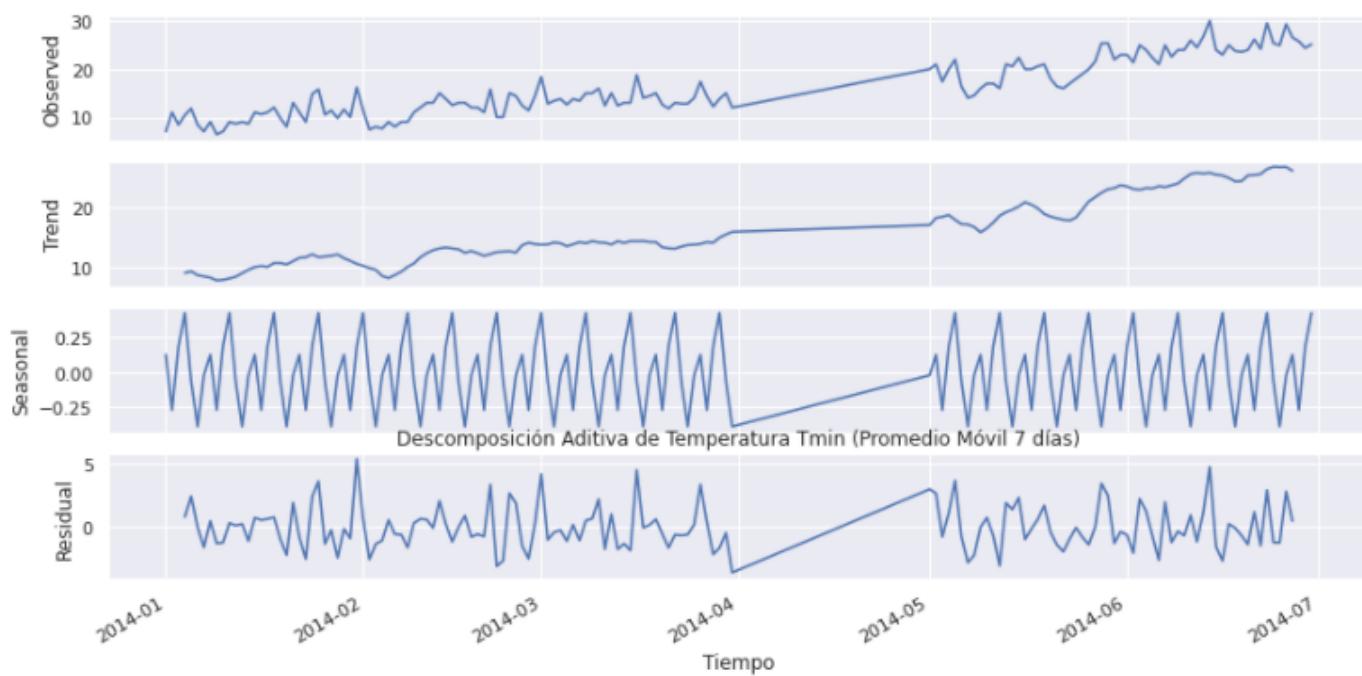
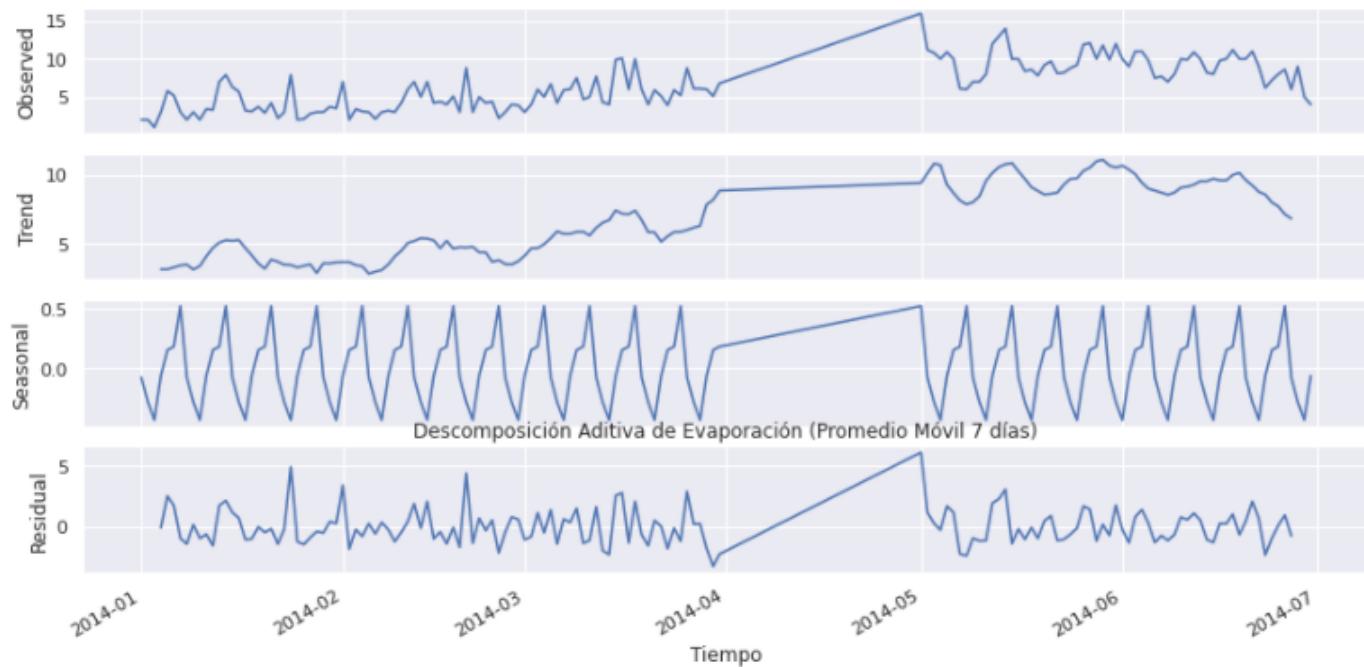
**Figura 20****Figura 21**

**Figura 22****Figura 23**

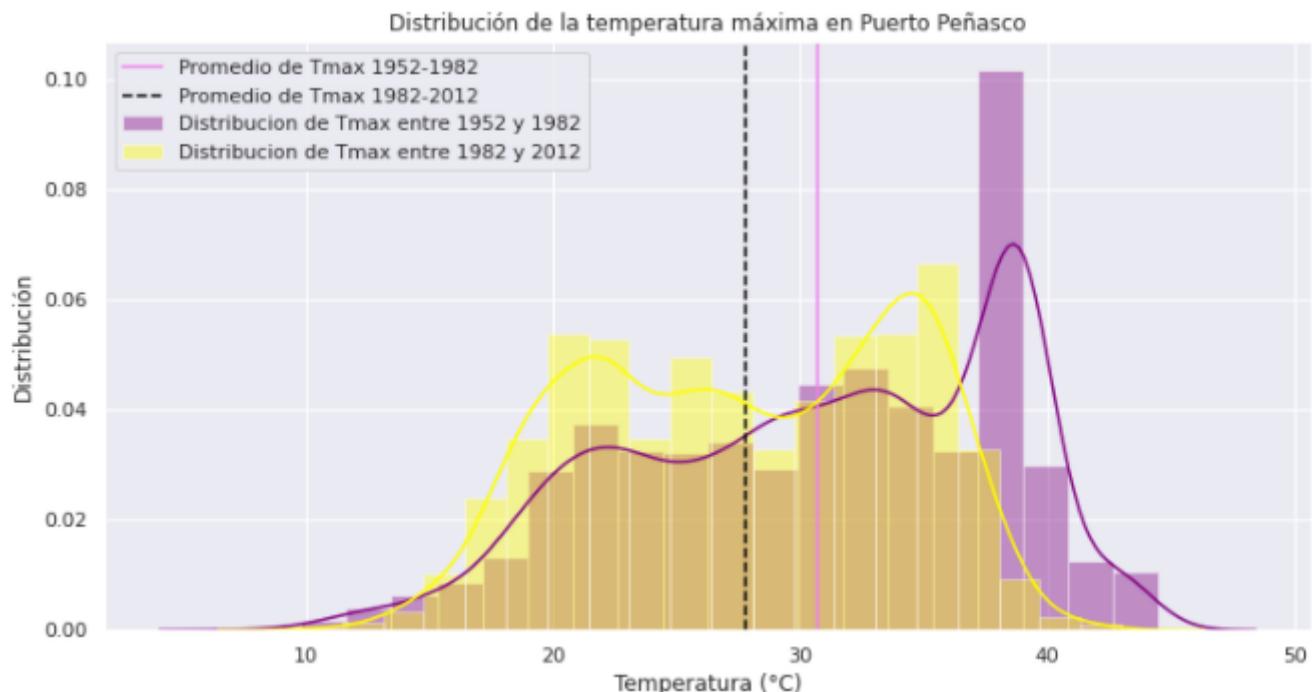
**Figura 24**

### 2.1.5. Actividad 5.5

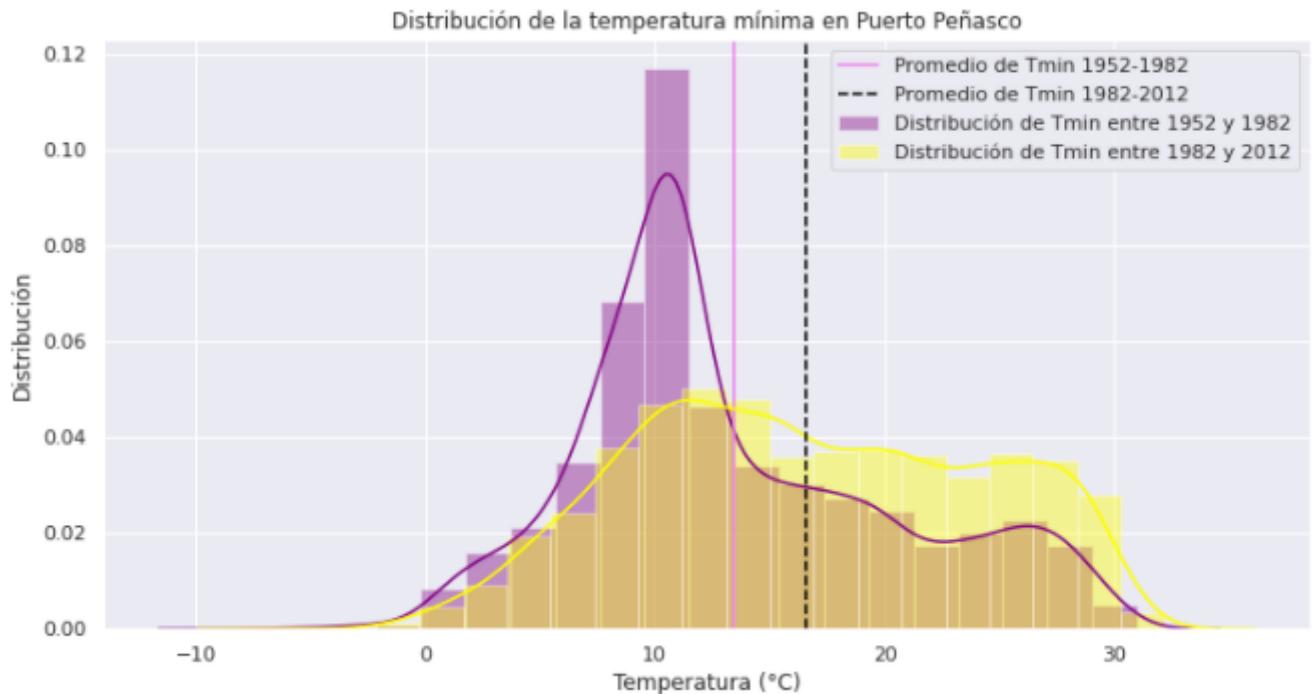
**Figura 25**

**Figura 26****Figura 27**

### 2.1.6. Actividad 5.6



**Figura 28**



**Figura 29**

### 3. Conclusión

En la primera actividad, se pudo observar que el comportamiento de las variables a través de los años es muy peculiar, con muchas fluctuaciones y muchas incoherencias, tal como que la temperatura parece mantenerse en los 40 grados por 3 años. También se puede observar que hay muchos lapsos en los que faltan datos, y en la variable de evaporación está casi completamente vacía.

En la actividad 5.2 podemos observar como la temperatura de los últimos 30 años se ha mantenido casi igual, con unos pequeños aumentos en los máximos de los últimos años. En cuanto al análisis de un año en específico (en esta ocasión se escogió el año 2010), podemos ver como tanto la Tmax como la Tmin aumentan en los meses de verano.

En cuanto a la actividad 5.3, donde se hizo lo mismo que en la 5.2, esta vez usando la variable de evaporación, podemos observar que la evaporación parece tener un patrón en la primera gráfica, de una manera senoidal, casi. También podemos observar que la segunda gráfica nos dice que en el 2010, la evaporación comienza a aumentar en abril y a disminuir en octubre, tal vez por las temperaturas más altas, lo que ocasionó mayor evaporación del mar. Por último, podemos ver que en la última gráfica la evaporación suele aumentar en los meses de verano, a la mitad del año, donde las distribuciones de las evaporaciones parecen ser casi uniformes por lo que nos enseñan las cajas.

En la actividad 5.4 podemos observar realmente cuanta falta de datos de evaporación tenemos, ya que cuando hacemos la descomposición de los datos de los últimos treinta años, no tenemos casi nada de datos. Las otras variables se grafican de una manera aceptable.

También, gracias a que algunas variables como precipitación son muy bajas (casi no llueve), la mejor forma para analizar esta variable es usando un promedio móvil de 7 días, ya que este nos enseña mejor las fluctuaciones de la misma.

En la actividad 5.5 se tuvo el mismo problema que se había mencionado antes; fue muy difícil encontrar un periodo en el que hubieran datos completos, por lo que decidí usar un periodo de 6 meses, cambiando la frecuencia a 7 días. Esto dio unas gráficas que muestran una anomalía que se generó entre abril y mayo del año en cuestión, donde se ve un incremento casi lineal de las variables. Esto debe de ser un error que se generó debido a algún problema en la estación, ya que es demasiado raro que hayan registrado datos tan extremos.

En la actividad 5.6 podemos observar como la diferencia entre periodos de 30 años contiguos no es mucho, sin embargo podemos ver que la temperatura en general en los treinta años más cercanos supera a la temperatura de los treinta pasados a esos, exceptuando unas anomalías en las temperaturas de alrededor de 38 y 10 grados, donde los primeros treinta años superan por mucho a los siguientes. Esto se puede deber a las anomalías que habíamos visto en las primeras gráficas.

Por otra parte, la actividad me pareció muy enriquecedora. Siento que aprendí mucho y me parece interesante el poder llevar a cabo un análisis de datos como este sin que se requiera mucha complejidad. Lo que más se me dificultó fue el encontrar intervalos de tiempo en los que no hay ningún dato nulo. Si pudiera agregarle algo a la actividad, sería, precisamente, una sección de código capaz de determinar dicho intervalo para que no tuviera que hacerse de manera manual y que no haya lugar a errores por visión. Esta actividad, en general, no me resultó tan complicada pero sí un poco extensa. En conjunto, le daría un nivel medio.

Además, tuve dificultades que me orillaron a realizar una entrega tardía ya que me quedé sin equipo de cómputo el día de entrega solicitada (se puede entregar evidencia de ello si así se solicita.).