

Introducción a la Biblioteca Pandas de Python

Fidel Alejandro Navarro Salazar

11 de febrero de 2019

1. Introducción

El objetivo de la actividad consistió en familiarizarnos con la biblioteca de Pandas, mediante la solución de problemas estadísticos utilizando datos meteorológicos.

2. Experiencia

2.1. Pandas en Python

La herramienta de Pandas puede llegar a ser muy útil ya que permite al usuario desarrollar análisis estadísticos sin la necesidad de escribir mucho código. Sin embargo, es necesario conocer e informarse de todas las funciones que Pandas ofrece, de lo contrario se puede llegar a complicar el uso de la biblioteca.

2.2. Análisis de Datos

Como parte de la actividad se realizó un análisis estadístico de los datos publicados por el Servicio Meteorológico Nacional sobre normales climatológicas. Para esta actividad se eligieron los datos publicados sobre Bahía de Kino, los cuales mostraban las temperaturas máximas y mínimas, la evaporación y la precipitación diaria durante los años 1974 al 2011.

2.3. Gráficas

Por medio del estudio de los diferentes parámetros fue posible realizar diferentes gráficas que muestran el comportamiento de la zona durante el año.

Como se puede visualizar en la Figura 1 los meses más lluviosos del año en la zona son Agosto y Septiembre, mientras que aquellos más secos son los de Abril y Mayo.

Por otro lado, en la Figura 2 podemos observar la precipitación por año desde 1974 al 2011. De esta gráfica no es posible realizar un análisis adecuado ya que existe una gran cantidad de años que no fueron observados. Sin embargo podemos observar que en comparación, los años más recientes han sufrido de una menor precipitación que en años posteriores.

En la Figura 3 podemos observar el comportamiento de la temperatura máxima a lo largo del año, y podemos observar que los meses más calurosos corresponden a los de Julio y Agosto, alcanzando temperaturas de más de 30°C.

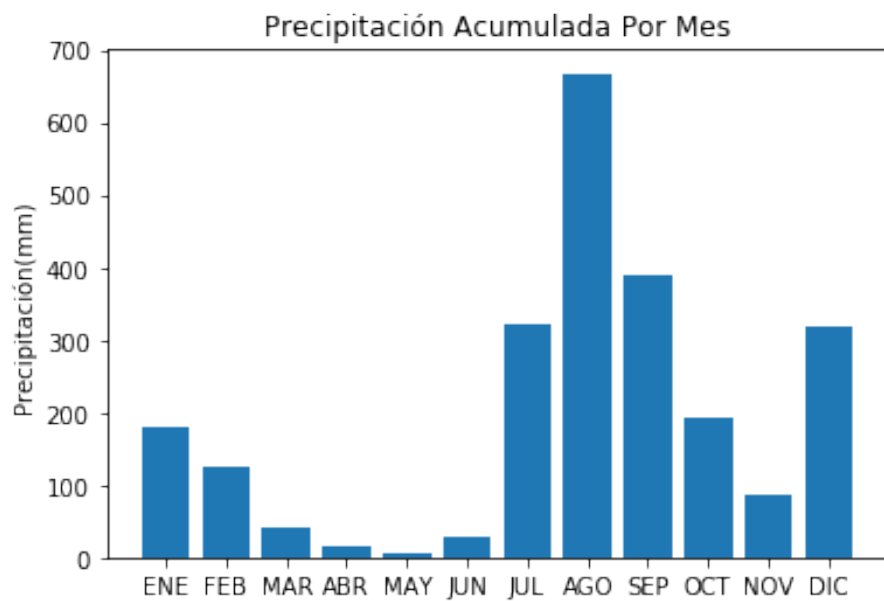


Figura 1: Gráfica de la precipitación acumulada por meses

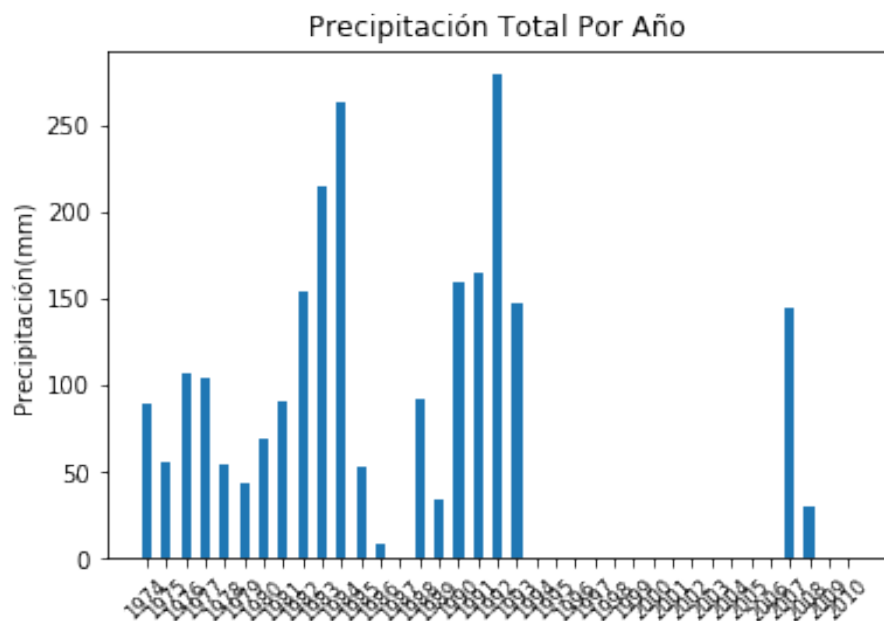


Figura 2: Gráfica de la precipitación acumulada por año.

De igual manera, en la Figura 4 podemos observar el comportamiento de la temperatura mínima registrada por año, donde los meses de Diciembre y Enero presentan las temperaturas más bajas que oscilan alrededor de los 5C.

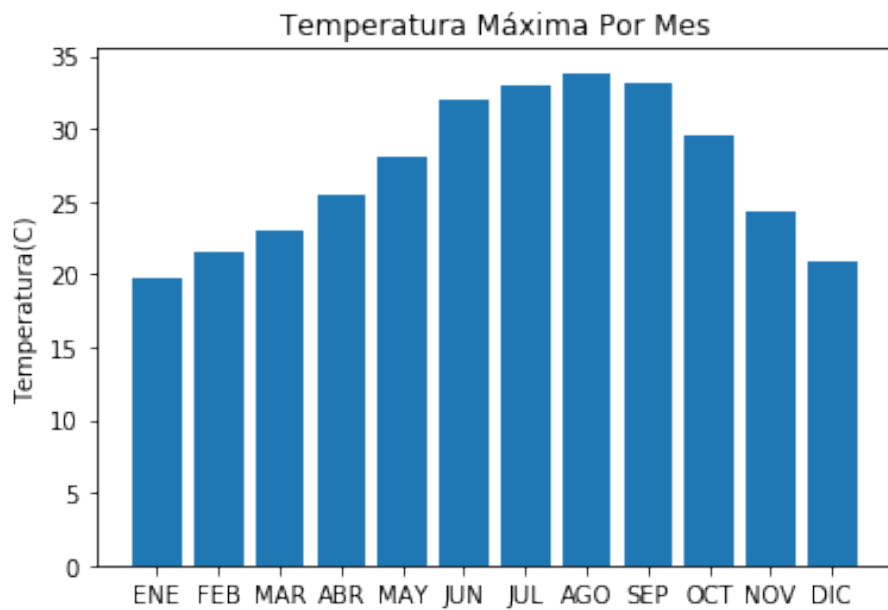


Figura 3: Gráfica de la temperatura máxima por mes

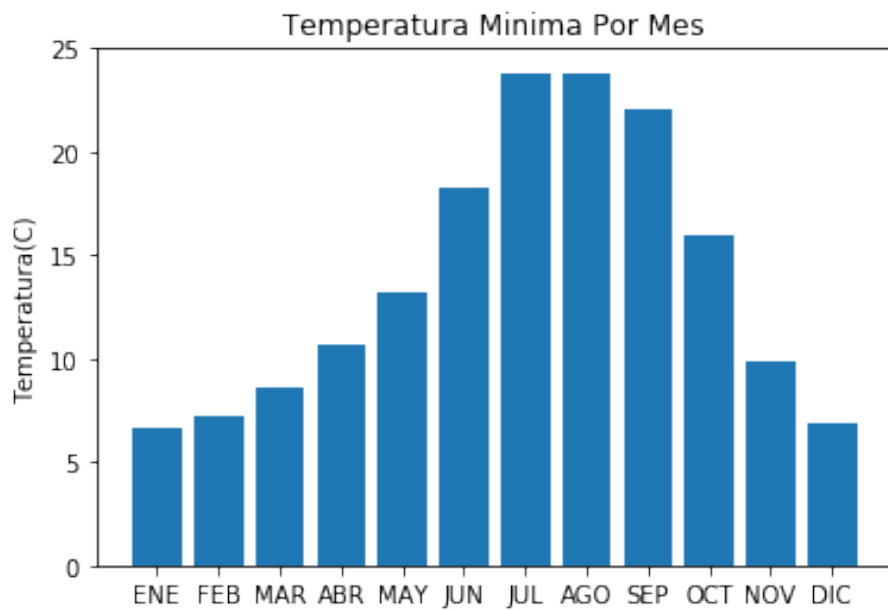


Figura 4: Gráfica de la temperatura mínima por mes.

2.4. Metodología

Para el desarrollo de las gráficas fué necesario reestructurar la información, por ello se creó una tabla de datos con los diferentes meses y se sumaron o

promediaron los valores correspondientes a cada mes, con la finalidad de facilitar la creación de gráficas.

3. Conclusión

Python y la biblioteca Pandas resultaron muy útiles para el análisis de grandes cantidades de datos, sin embargo fue necesario realizar una busque intensiva de fuentes para el uso correcto de la biblioteca debido a que no se tenía conocimiento previo de varias funciones que se requerían para el manejo adecuado de los datos.