

Reporte de la Actividad 4

Marco Antonio Cabello López

Grupo 1

Domingo 17 de Febrero del 2019

1 Introducción

Primeramente y como introducción, en esta actividad nos enfocamos en el uso de la biblioteca Matplotlib, en un entorno de Jupyter Notebooks y en el contexto de análisis de datos. El análisis se realizó sobre los datos meteorológicos de Valle de Bravo, un municipio del estado de México, los cuales fueron proporcionados por el Servicio Meteorológico Nacional.

En esta actividad iniciaremos con las tareas de visualización de gráficos estadísticos con los datos diarios de precipitación, temperaturas mínimas y máximas proporcionados por el Servicio Meteorológico Nacional.

2 Desarrollo de la actividad

2.1 Metodología

Se importaron las siguientes bibliotecas en el archivo de Jupyter Notebooks para trabajar con los datos de Valle de Bravo:

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

Estas nos ayudarán con los cálculos matemáticos, el análisis estadístico y la realización de gráficas. Después se leyó el archivo mediante el comando “pd.read_csv”, y posteriormente se creó un Data Frame mediante el comando “pd.DataFrame()”. Como el archivo cuenta con una columna de fechas, a esta se le dio formato de fecha mediante:

```
df['FECHA'] = pd.to_datetime(df.apply(lambda x: x['FECHA'], 1)
```

Table 1: Funciones de Data Frame empleadas.

Función	Acción
df.head()	Encabezado
df.tail()	Final
df.dtypes()	Tipos de variables
df.mean()	Promedio
df.std()	Desviación Estándar
df.median()	Mediana
df.max()	Máximo
df.min()	Mínimo
df.describe()	Resumen estadístico
df.sum()	Sumar
df.shape()	Estructurar Filas y Columnas
df.index()	Índice
df.count()	Conteo de datos
df.corr()	Correlación
df.PRECIP.max()	Precipitación Máxima
df.unique()	Valores diferentes

Para elaborar las gráficas se emplearon diversas funciones sobre el Data Frame. Determinamos el número de años que hay en el archivo mediante el conteo de los datos, obtuvimos 54 años, los cuales se asignaron con el comando:

```
init = 1961
AÑOS = [init + i for i in range(0, 54)]
PAÑO = [ df[df.AÑO==(init + i)].PRECIP.sum() for i in range(0, 54)]
```

Para tener arreglos con días, meses y años declaramos el de los meses de 0 a 12, y el de los años de 1961 hasta 2014. Para tener los nombres de los meses, también se creó otro arreglo:

```
df['MES'] = df['FECHA'].dt.month
df['AÑO'] = df['FECHA'].dt.year
```

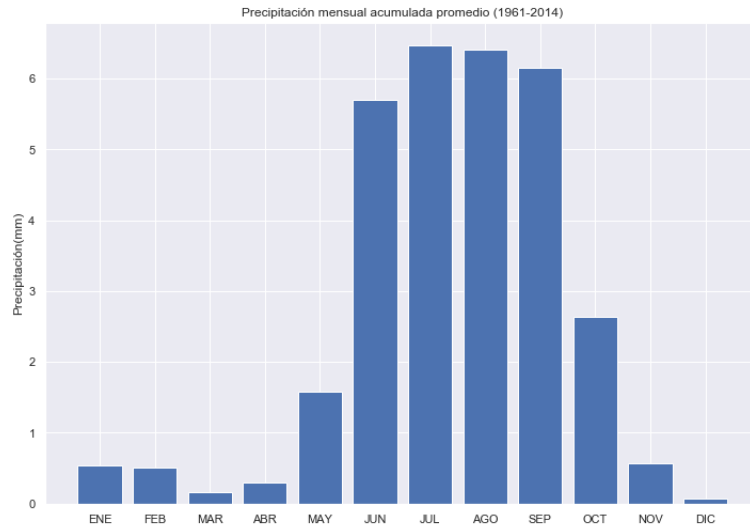
Para hacer los gráficos de caja se recurrió a la biblioteca Seaborn, ya que no fue posible realizarla mediante Matplotlib. Recurrimos a los siguientes comandos para hacerlo a partir del DataFrame:

```
ax = sns.boxplot(x="AÑO", y="TMIN", data=df)
plt.xticks(rotation=45)
plt.title("Temperatura mínima promedio anual para cada año (1961-2014)")
plt.ylabel ("Temp C")
```

2.2 Resultados

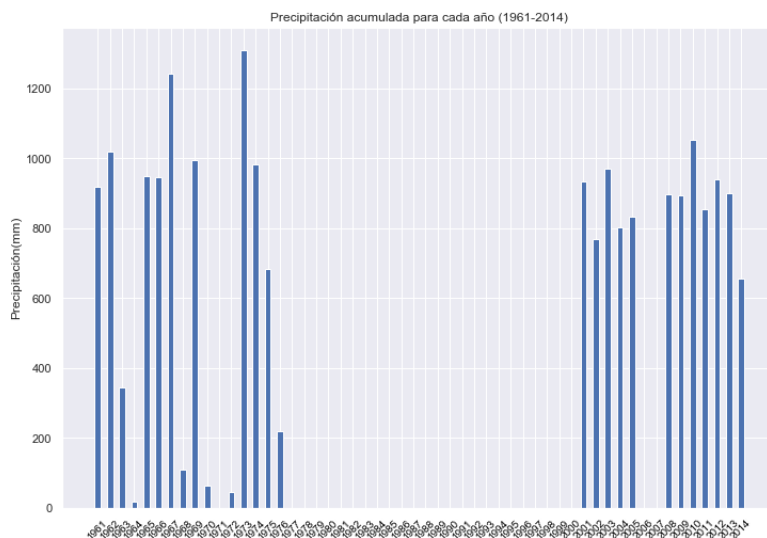
1. Elabora una gráfica de barras (barplot) de precipitación mensual acumulada promedio de la colección de datos de la estación que se esté analizando.

Podemos apreciar que los meses más lluviosos han sido Julio, Agosto y Septiembre.



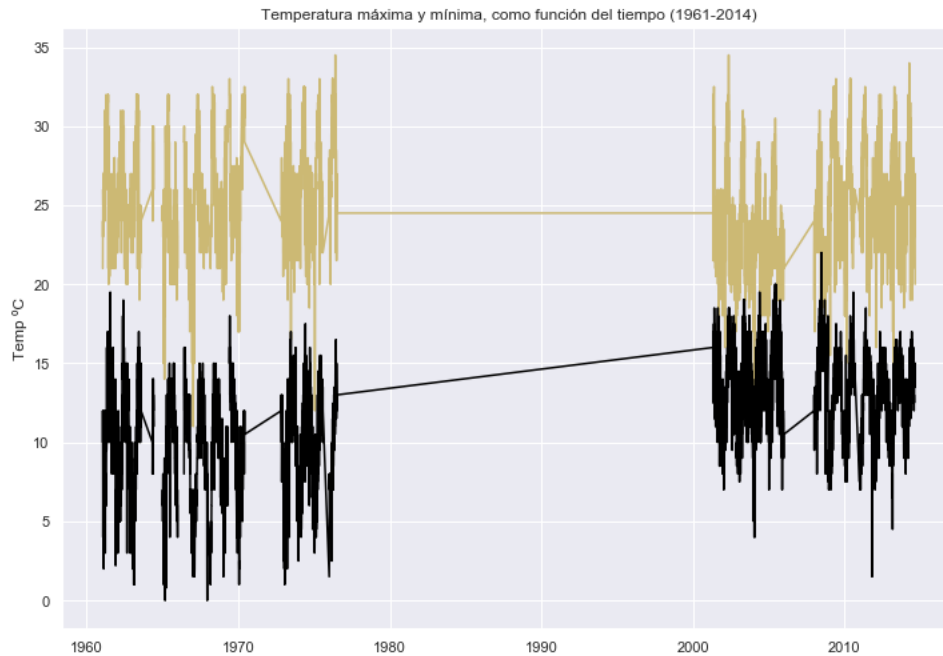
2. Elabora una gráfica de barras de precipitación acumulada para cada año de la misma colección.

Podemos afirmar que los años con mayores precipitaciones fueron 1957, 1973 y 2010, y que el archivo carece de datos de 1977-2000.



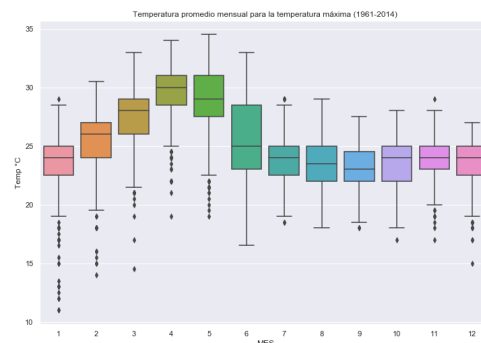
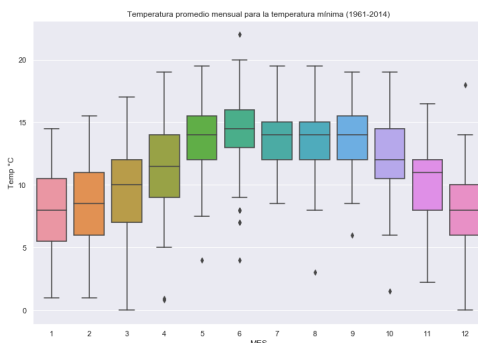
3. Elabora una gráfica de la evolución de la temperatura máxima y mínima en la misma figura, como función del tiempo de la colección de datos.

Determinamos la evolución temporal de la temperatura máxima (amarillo) y mínima (negro) en el municipio de Valle de Bravo, notamos que la temperatura ha variado en un rango de aproximadamente 0-35°C.



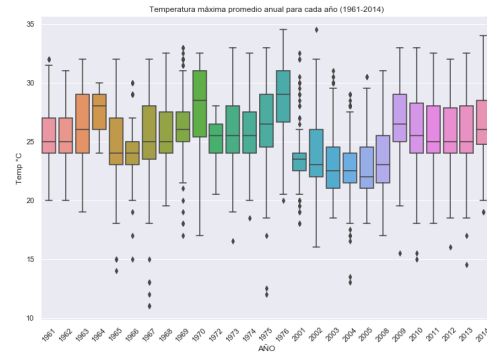
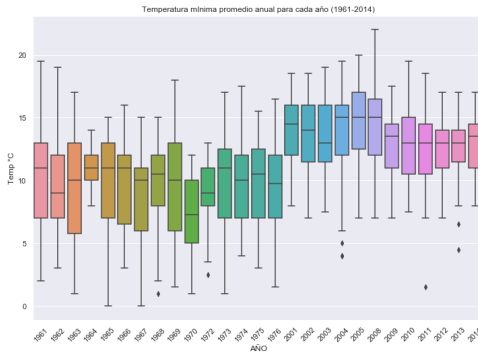
4. Elabora una gráfica de cajas (boxplot) de la temperatura promedio mensual para la temperatura mínima y máxima por separado.

Podemos observar que los meses de Enero, Febrero y Diciembre presentaron la menor temperatura mínima promedio, y los meses de Abril y Mayo presentaron la mayor temperatura máxima promedio.



5. Elabora una gráfica de cajas de la temperatura mínima y máxima promedio anual para cada año por separado.

Podemos analizar que el año de 1970 presento la menor temperatura mínima promedio y que de 2001-2014 la temperatura mínima promedio es mayor en comparación con la temperatura mínima promedio de la época 1961-1976, y los meses de Abril y Mayo presentaron la mayor temperatura máxima promedio, además el año de 1970 presento la mayor temperatura máxima promedio.



3 Conclusiones

Podemos concluir que la biblioteca Matplotlib y Seaborn son herramientas muy útiles para visualizar datos y realizar un análisis de los mismos a partir de las gráficas, observamos diversos aspectos interesantes del municipio de Valle de Bravo con la información que nos proporciona la correspondiente estación del Servicio Meteorológico Nacional.

4 Referencias

- El Servicio Meteorológico Nacional, Información Estadística Climatológica, Recuperado de:
<https://smn.cna.gob.mx/es/climatologia/informacion-climatologica/informacion-estadistica-climatologica>
- Matplotlib. Recuperado de:
<https://matplotlib.org/>
- Seaborn: statistical data visualization. Recuperado de:
<https://seaborn.pydata.org/generated/seaborn.boxplot.html>