**Proyecto de Scraping de Datos Climáticos**

1. **Proceso de extracción, tratamiento y preparación de datos**

**Descripción del proceso**

Los datos climáticos fueron obtenidos del sitio web Climate at a Glance Global Time Series [[1]](#footnote-1) de la Oficina Nacional de Administración Oceánica y Atmosférica (NOAA, siglas en inglés) a través de web scraping utilizando la librería Selenium y BeautifulSoup en Python. Se diseñaron dos scripts principales para esta tarea:

1. *scraper\_clima.py*: Este script se encarga de generar las URLs necesarias y de realizar el scraping de los datos climáticos.
   * Función *get\_climate\_url(parameter, first\_year, last\_year)*: Genera la URL específica para acceder a los datos climáticos de la NOAA. Los parámetros incluyen el tipo de datos (‘tavg’ para anomalías de temperatura y ‘pcp’ para precipitación) y el rango de años, por defecto de 1979 a 2024 (última información disponible).
   * Función *scrape\_climate\_data(climate\_url)*: Utiliza Selenium para cargar la página web y BeautifulSoup para extraer el contenido de la tabla HTML que contiene los datos climáticos a partir de la URL generada con la función anterior.
2. *run\_scraper.py*: Este script se encarga de realizar el scraping de datos climáticos (llamando a las funciones del script *scraper\_clima.py*), procesar los datos extraídos, renombrar columnas, convertir las columnas a formatos adecuados y guardar los datos en archivos CSV.
   * Función *procesar\_datos(climate\_url, parameter)*: Realiza el scraping utilizando las funciones de *scraper\_clima.py*, renombra las columnas, convierte las fechas y guarda los datos en archivos CSV para su uso posterior en el dashboard de visualización de Streamlit.
3. **Dashboard y visualizaciones**

**Descripción del dashboard:**

Este dashboard interactivo realizado con Streamlit y Plotly, presenta una visualización detallada de las anomalías globales en temperatura y precipitación desde 1979 hasta 2024 (aunque permite modificar las fechas de análisis). A través de gráficos dinámicos y filtros personalizables, permite explorar cómo han variado las condiciones climáticas a lo largo del tiempo.

**Características principales:**

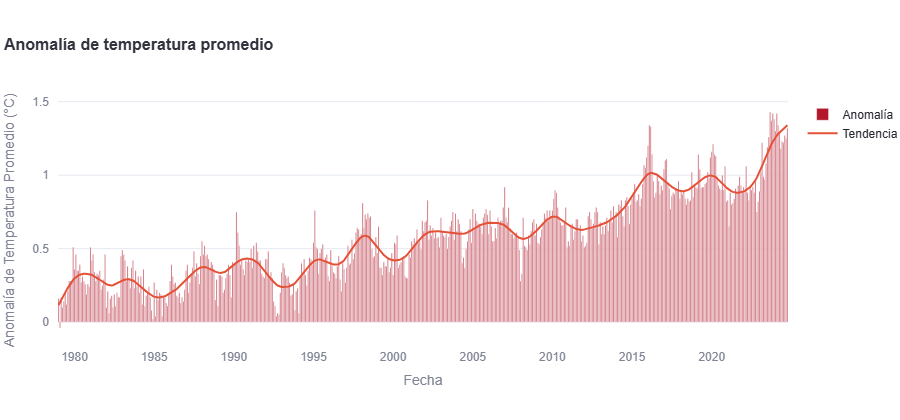
* **Anomalías de Temperatura Promedio (TAVG):** Gráfico de barras con línea de tendencia que muestran las desviaciones de la temperatura media en comparación con los valores históricos.
* **Anomalías de Precipitación (PCP):** Visualización de las variaciones en las precipitaciones acumuladas en milímetros a lo largo de los años.
* **Rango Temporal Personalizable:** Utiliza los filtros de fecha para ajustar el período de tiempo que deseas analizar.
* **Datos Ordenados y Destacados:** Visualiza los 10 meses con mayores anomalías en temperatura y precipitación, ordenados y presentados para una fácil interpretación.



**Descripción de las visualizaciones:**

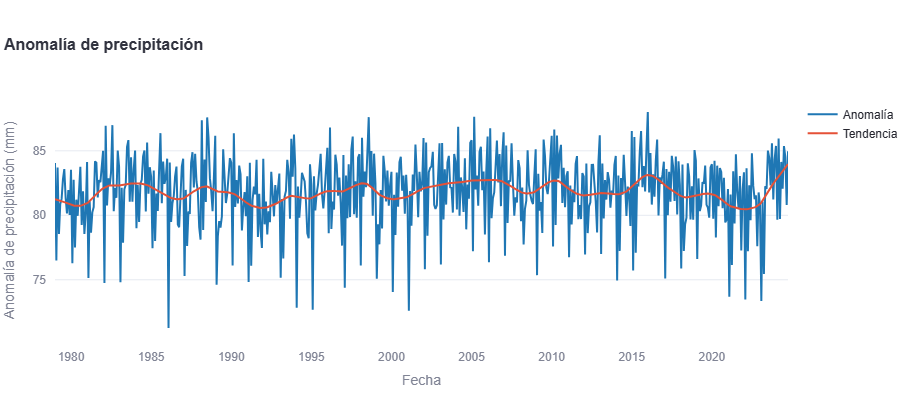
Las visualizaciones fueron diseñadas para presentar de manera clara y comprensible las anomalías en temperatura promedio y precipitación a lo largo del tiempo. Se utilizaron dos gráficos principales:

1. Gráfico de barras con tendencia para las anomalías de temperatura:



* + Objetivo: Mostrar la variación de la anomalía de temperatura promedio a lo largo del tiempo, así como la línea de tendencia a lo largo del tiempo.
  + Justificación: Las barras permiten ver fácilmente las diferencias entre cada mes, resaltando períodos con anomalías significativas.

1. Gráfico de líneas con tendencia para anomalías de precipitación:



* + Objetivo: Mostrar la tendencia en la anomalía de precipitación a lo largo del tiempo y evidenciar la naturaleza cíclica de las precipitaciones.
  + Justificación: Las líneas son ideales para visualizar tendencias y patrones a lo largo del tiempo, permitiendo detectar períodos de aumento o disminución en la precipitación.

1. **Insights obtenidos**

Anomalías de Temperatura

* Se visualiza una clara tendencia positiva en el incremento de las anomalías de temperatura promedio a medida que transcurre el tiempo, esta tendencia se hace más notable a partir de 1990.
* En los últimos 10 años se puede ver que estas anomalías aumentan a un ritmo más acelerado. Asimismo, se identifica que 8 de las 10 anomalías de temperatura más altas desde 1979 se dieron en los últimos 2 años.

Anomalías de Temperatura

* En este gráfico se puede observar la naturaleza cíclica de las precipitaciones. Así, por ejemplo, se identifica al mes de febrero como el mes con menores anomalías, mientras que marzo es uno de los meses con mayores anomalías.
* Por otra parte, se visualizar una ligera tendencia positiva en las anomalías de precipitaciones, por lo que es muy probable que estas anomalías se incrementen en los próximos años.

1. Disponible en: <https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/global/time-series/globe/tavg/> [↑](#footnote-ref-1)