

I) Resumen Ejecutivo

El presente proyecto tiene como objetivo analizar las anomalías de temperatura globales a lo largo de los años, utilizando un conjunto de datos históricos obtenidos de la NASA, y desarrollar un modelo predictivo para anticipar posibles tendencias futuras. Además, se busca crear un dashboard interactivo para visualizar los insights clave de forma intuitiva y accesible.

Se realizaron varias etapas de limpieza y exploración de los datos, así como la implementación de múltiples técnicas de machine learning, incluyendo Random Forest y regresión polinómica, para obtener predicciones precisas.

II) Objetivo

- Analizar la evolución de las anomalías de temperatura global desde 1880.
- Identificar patrones significativos y factores que influyen en los cambios de temperatura.
- Predecir tendencias futuras utilizando técnicas de machine learning.
- Proporcionar una herramienta visual para comunicar los hallazgos y predicciones de manera interactiva.

III) Preguntas e Hipótesis

1. ¿Existen patrones significativos en las anomalías de temperatura por década?
2. ¿Cuáles son los años con anomalías extremas y qué los caracteriza?
3. ¿Es posible modelar con precisión las tendencias de temperatura futura utilizando datos históricos?

Hipótesis: La tendencia general de las anomalías de temperatura es al alza debido al cambio climático antropogénico, y los modelos avanzados pueden proporcionar predicciones confiables.

IV) Contexto Comercial

Comprender y predecir las tendencias de temperatura es fundamental para gobiernos, empresas y organizaciones que buscan mitigar los efectos del cambio climático. Este proyecto proporciona un modelo que puede ayudar a identificar regiones vulnerables y planificar estrategias sostenibles.

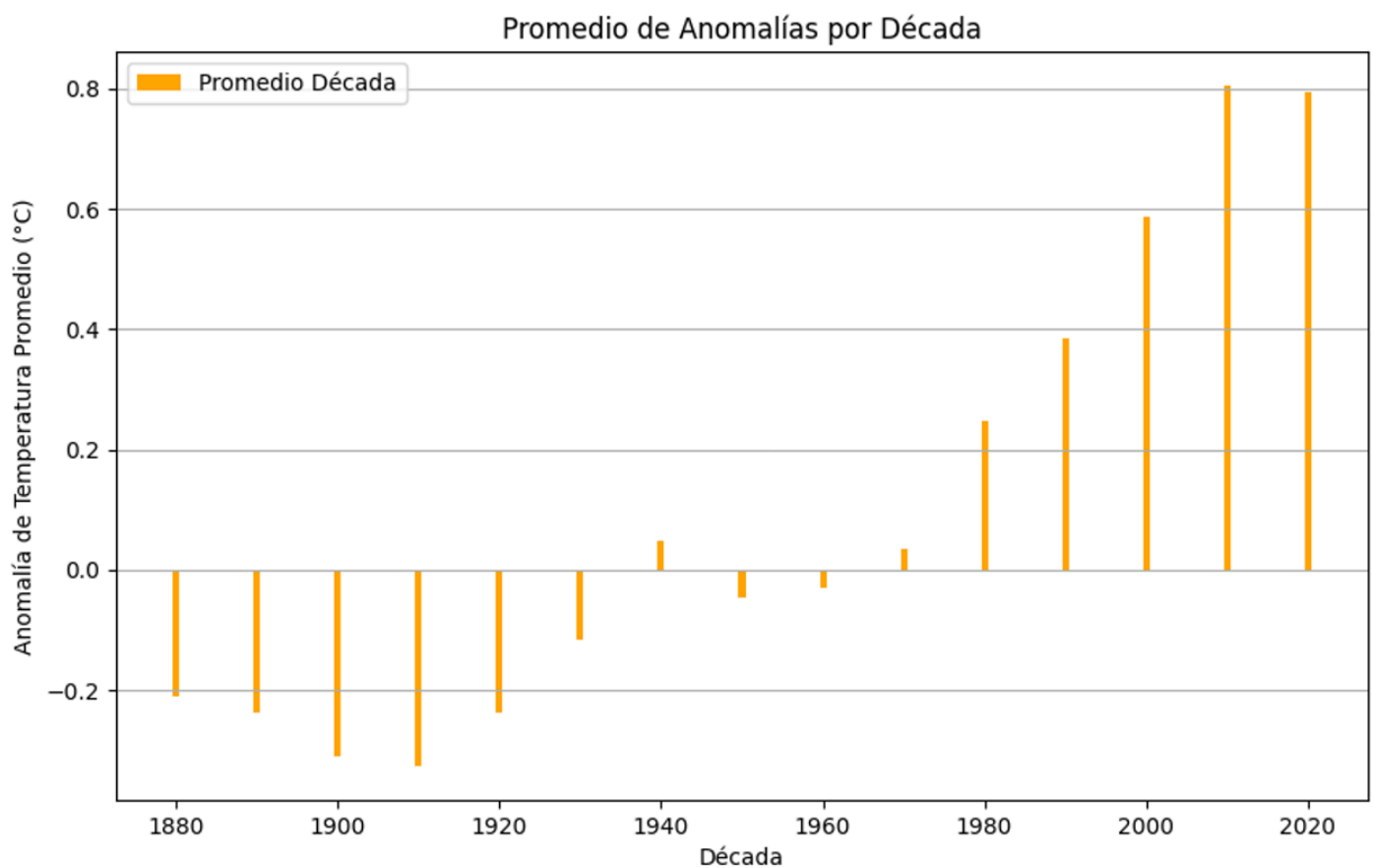
V) Problema Comercial

La creciente frecuencia de eventos climáticos extremos representa un desafío para la planificación y toma de decisiones. Este proyecto busca ofrecer insights útiles y predicciones para anticipar estos cambios y reducir su impacto.

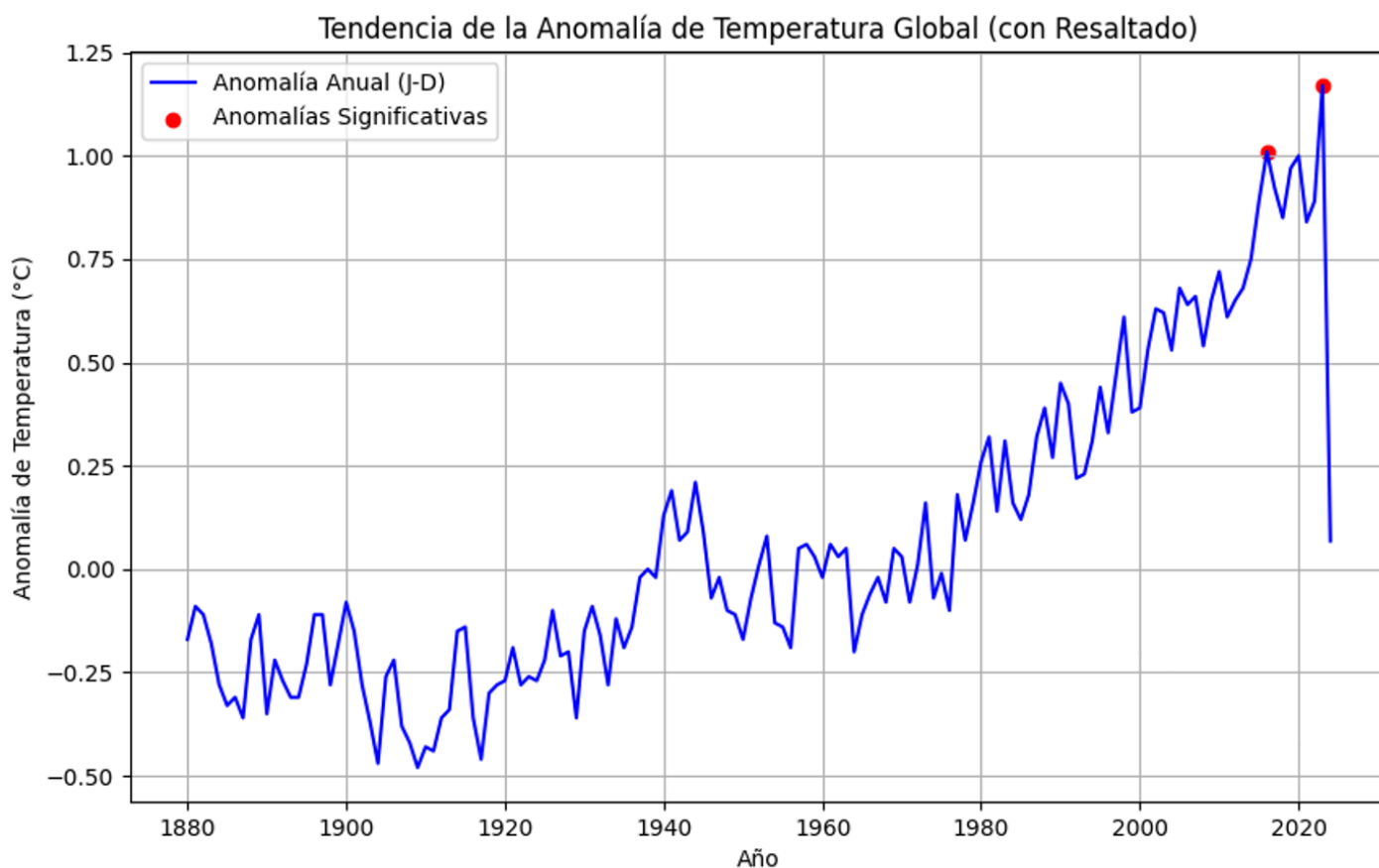
VI) Análisis Exploratorio de Datos (EDA)

Visualizaciones Incluidas:

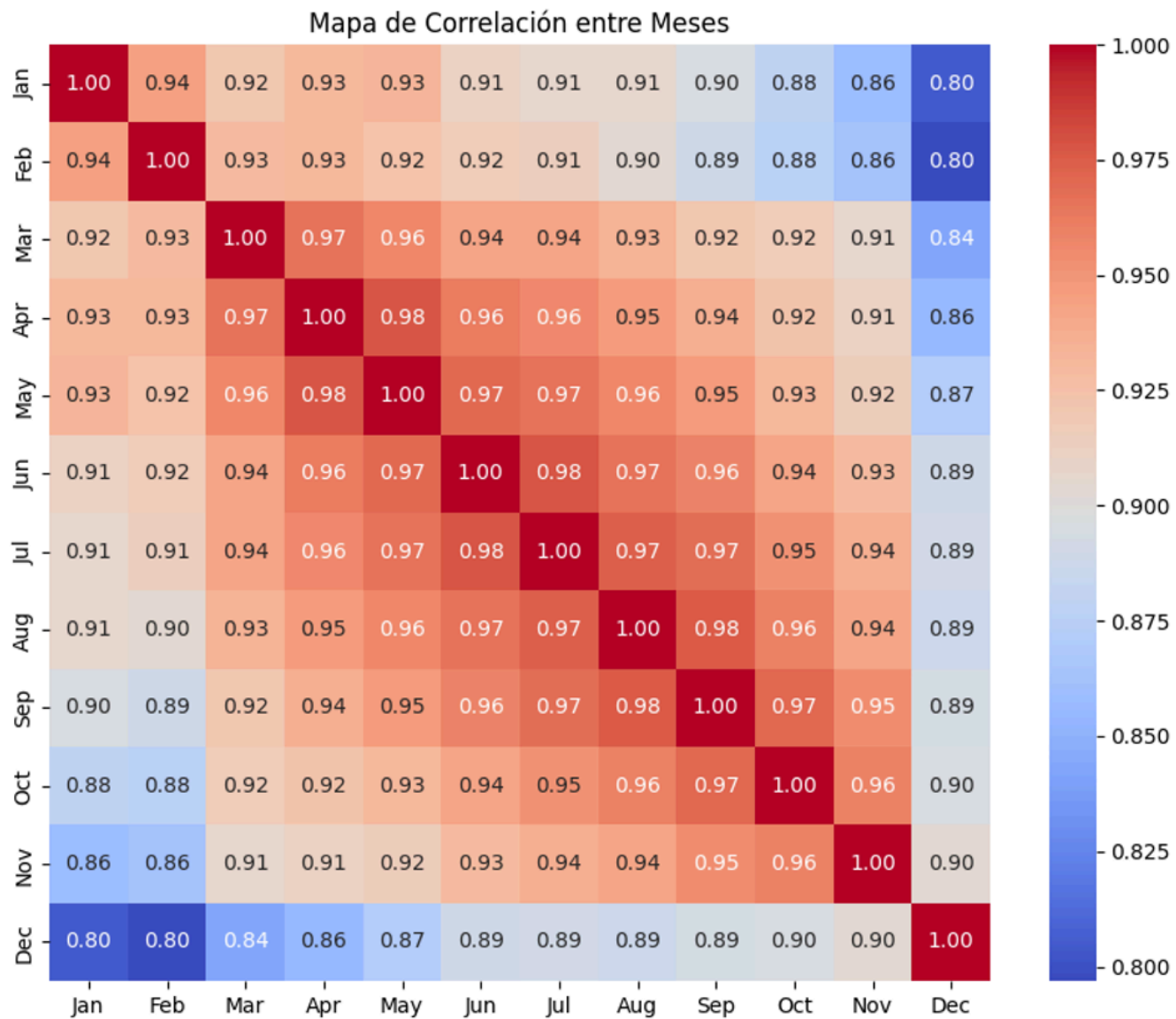
1. **Promedios por década:** Gráfico de barras que muestra las anomalías promedio por cada década.



2. **Cambios año a año:** Gráfico lineal que destaca los cambios en las anomalías de un año a otro.



3. **Tendencias por estación:** Gráfico que analiza las anomalías promedio por estación (DJF, MAM, JJA, SON).



Insights Clave del EDA:

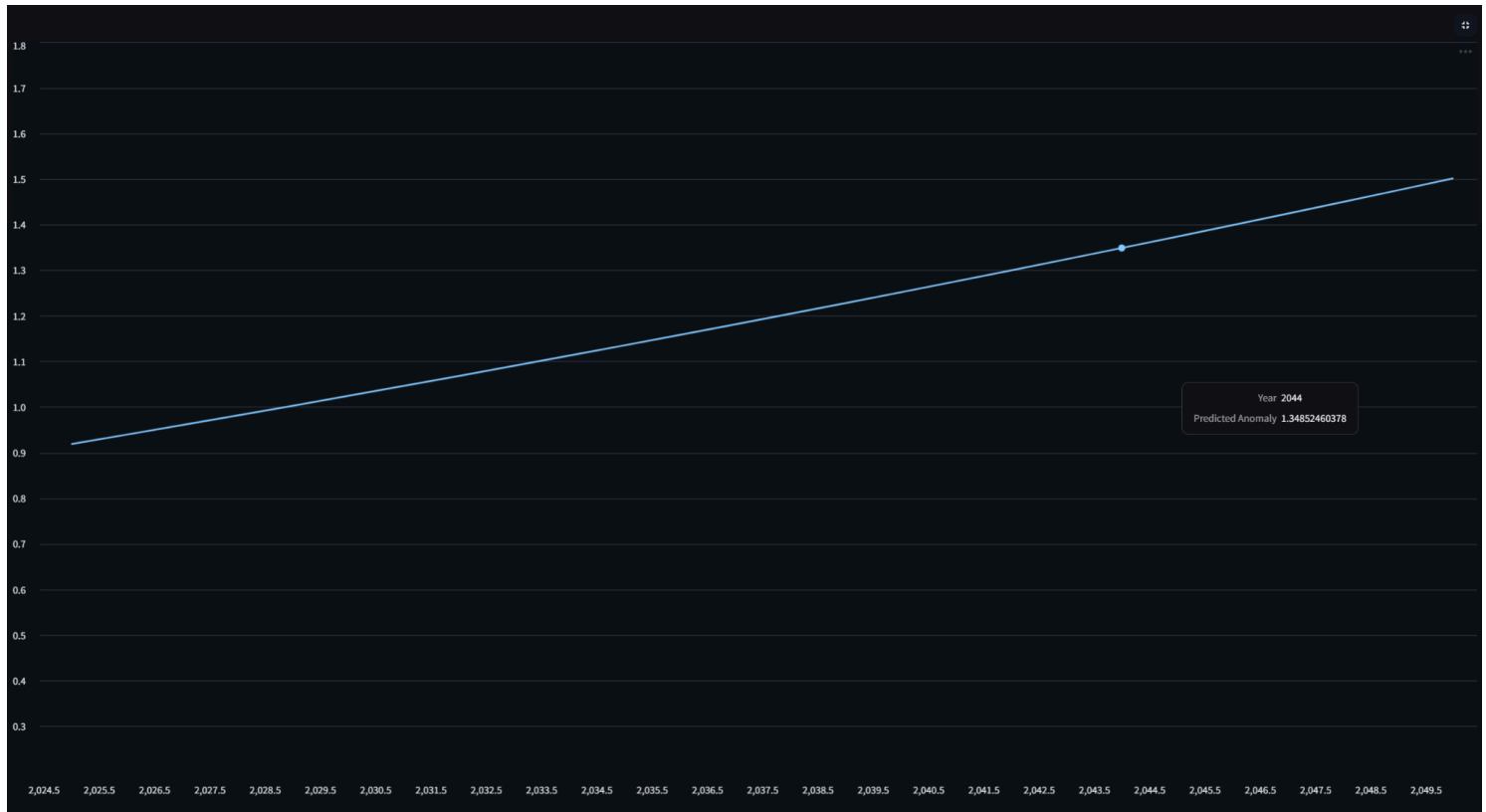
- Las anomalías de temperatura han mostrado un aumento constante desde el siglo XX.
- Existen patrones significativos por estaciones, siendo DJF la más impactada.

VII) Modelos de Machine Learning

Modelos Utilizados:

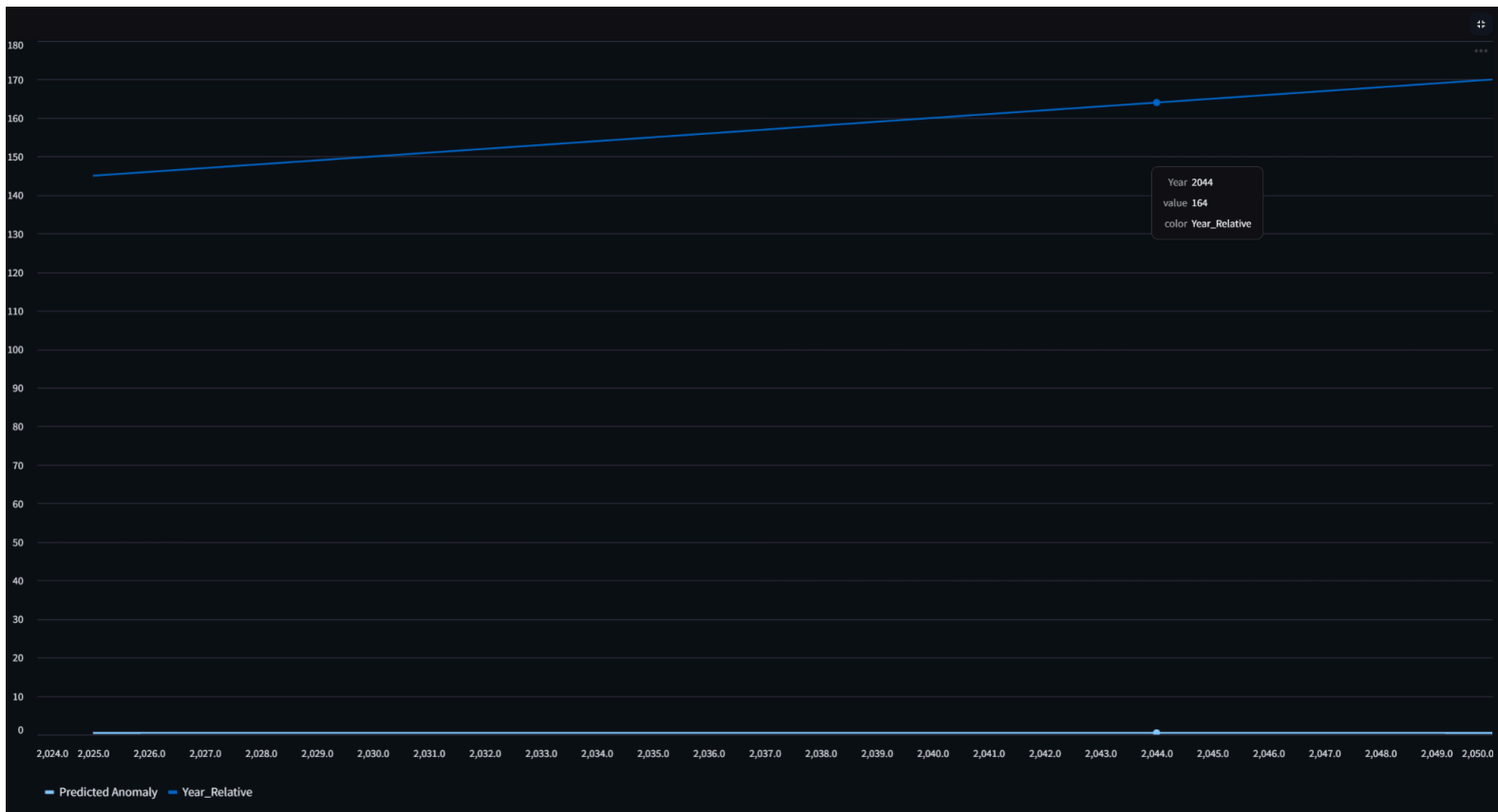
1. Regresión Polinómica:

- Se utilizó un modelo de grado 3 para capturar relaciones no lineales.
- **Resultados:** MSE de 0.0139.



2. Random Forest Regressor:

- Modelo robusto para capturar relaciones complejas en los datos.
- **Resultados:** MSE de 0.0099.



Comparación de Tecnologías Utilizadas:

- **Lenguajes y Librerías:**
 - Python (pandas, scikit-learn, Streamlit, Plotly).
- **Bases de Datos:**
 - PostgreSQL con conexión mediante SQLAlchemy.

Comparación de Modelos:

- La regresión polinómica mostró tendencias claras pero menos precisas en anomalías extremas.
- Random Forest proporcionó mejores resultados en general.

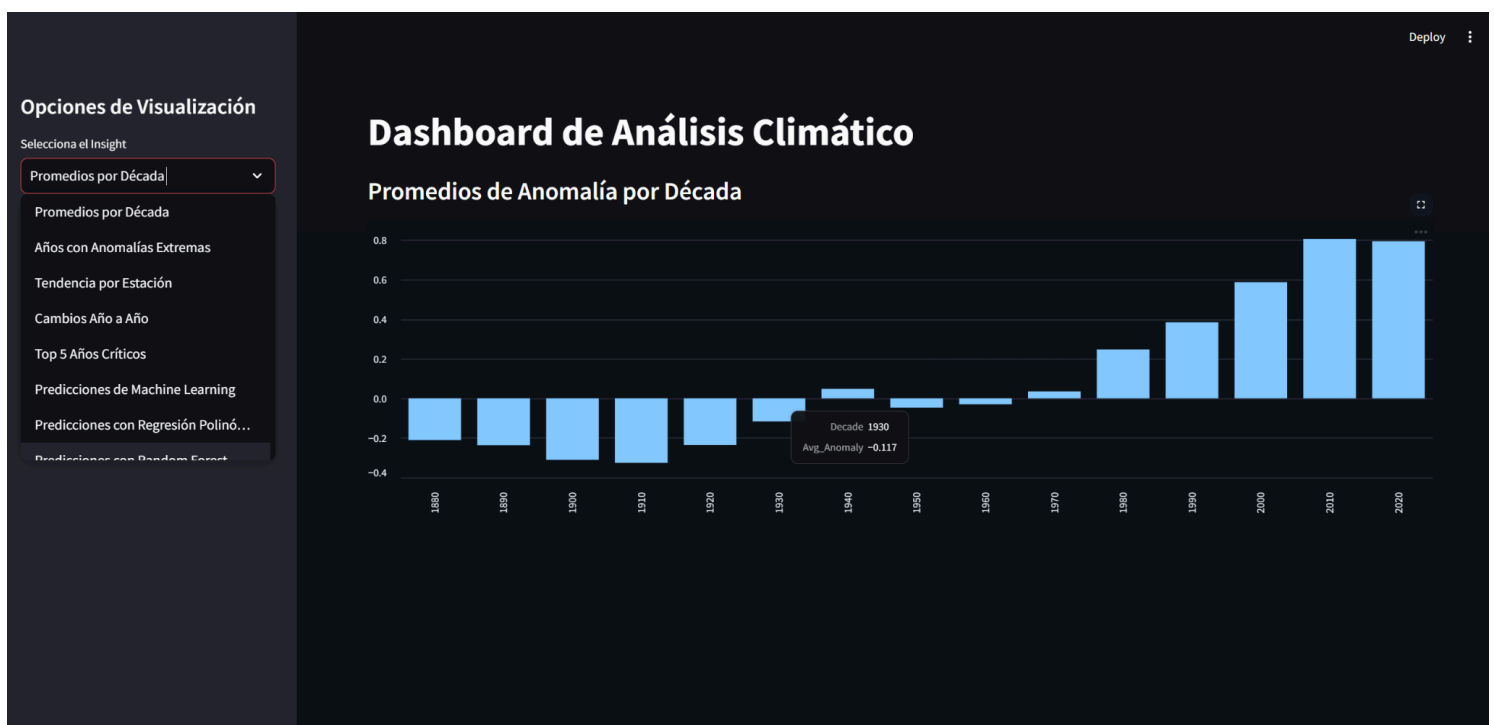
VIII) Insights Clave

1. Las anomalías de temperatura muestran una tendencia al alza preocupante.
2. Los modelos predicen un aumento significativo para las próximas décadas.
3. La combinación de visualizaciones y predicciones ofrece una herramienta valiosa para la toma de decisiones.

IX) Dashboard Interactivo

El dashboard creado para este proyecto se desarrolló utilizando Streamlit y permite a los usuarios explorar los datos y los insights clave de manera interactiva. Las funcionalidades principales incluyen:

- **Visualización de Datos:**
 - Promedios por década.
 - Cambios año a año.
 - Tendencias por estación.
- **Modelos Predictivos:**
 - Predicciones futuras utilizando Random Forest y regresión polinómica.
- **Interactividad:**
 - Navegación sencilla entre distintas visualizaciones.
 - Gráficos dinámicos que responden a las selecciones del usuario.



X) Limitaciones y Trabajo Futuro

Limitaciones:

- Los modelos no consideran factores externos como políticas climáticas o eventos naturales.
- Los datos históricos podrían no reflejar adecuadamente los cambios futuros.

Trabajo Futuro:

- Integrar modelos más avanzados como redes neuronales.
- Ampliar el conjunto de datos para incluir factores adicionales como emisiones de carbono y niveles del mar.

XI) Conclusión

Este proyecto demuestra el valor del análisis de datos y las técnicas de machine learning para abordar problemas complejos como el cambio climático. Los insights obtenidos y el dashboard interactivo proporcionan una base sólida para futuras investigaciones y acciones concretas.