

Assignment: Regional Climate Anomaly Trends Through Data Science

Andrés Navarro

1 Introduction

This assignment focuses on the analysis of regional climate anomaly trends from 1950 to 2024, using real-world climate data. Students will select a region of interest—such as a city, or region—and investigate how temperature anomalies have evolved over time.

The core objective is to compute monthly temperature anomalies by comparing observed values to a climatological reference period (1981–2010). This allows for the identification of deviations from typical climate conditions and the detection of long-term trends.

To support this analysis, students will work with the ERA5 reanalysis dataset, a high-resolution global climate dataset widely used in climate research. Through this project, students will apply key data science skills—including data wrangling, statistical analysis, and visualization—while gaining experience with climate data formats such as NetCDF.

This case study is designed to strengthen students’ ability to conduct exploratory data analysis, interpret scientific data, and communicate findings effectively—core competencies in the field of data science.

2 Objectives

- Explore and understand the structure of climate data in NetCDF format.
- Understand the concept of **climate anomaly**, including its scientific significance, and apply it effectively to analyze temperature data.
- Visualize anomalies through **time series plots** and **spatial maps**.
- Identify and interpret **regional trends and patterns** in temperature anomalies.
- Develop and communicate insights using data science tools and techniques.

3 Dataset Description

The ERA5 dataset is fully available at <https://cds.climate.copernicus.eu/>. An alternative NetCDF file named `ERA5-MONTHLY-TS-1940-2025.nc`, which contains monthly mean temperature data (`t2m`) from 1940 to 2025. The dataset covers the entire globe with a spatial resolution of 0.25 degrees.

4 Workflow

- Load and explore the provided NetCDF dataset.
- Calculate the climatological mean for the reference period (1981–2010).
- Compute the temperature anomalies for the period 1950–2024.
- Visualize the anomalies using time series and spatial maps.
- Analyze the results to identify significant patterns and trends.

5 Suggested Tools

- Climate Data Operators (CDO)
- Python libraries: `xarray`, `numpy`, `matplotlib`, `cartopy`

6 Deliverables

- Python and Bash scripts used for data processing and analysis.
- Visualizations, including time series plots and global or regional anomaly maps.
- A concise report summarizing the methodology, key findings, and interpretations.

7 Optional Extensions

- Evaluate alternative climatological baseline periods (1961-1990, 1951-1980)
- Compare anomalies with another dataset (e.g., CRUTS).
- Apply smoothing techniques or trend analysis.
- Investigate the impact of specific climate events on the anomalies.

8 Evaluation Criteria

The assignment will be evaluated based on the following aspects:

- **Correctness and completeness of the data analysis (40%):** Proper calculation of anomalies, use of reference period, and handling of the dataset.
- **Code quality and reproducibility (20%):** Well-structured, commented, and functional scripts.
- **Clarity and insight in visualizations (20%):** Effective use of plots to communicate trends and patterns.
- **Quality of the written report (20%):** Clear explanation of methods, results, and conclusions.