Data Mining Individual Project - Final Presentation
Advanced Data Mining - CS5803,
Master of Data Science and AI - CSE, UOM.

Detecting Climate Anomalies in Sri Lanka: A Data-Driven Approach

An Analysis of 13.5 Years of Weather Data Using Machine Learning

Deeghayu Adhikari - 258722D

20.07.2025

The Problem

Pinpointing Climate Instability

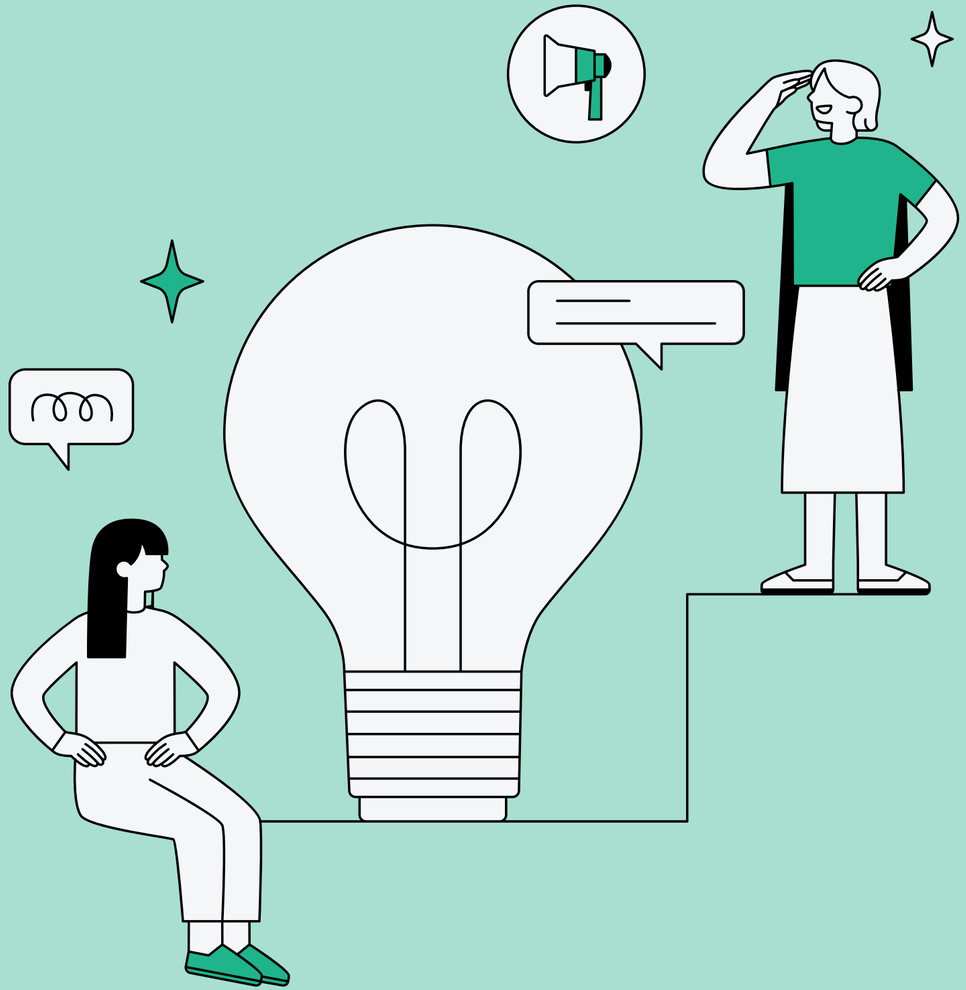
01. Global climate change manifests as local, extreme weather events.

02. Sri Lanka's diverse geography makes it vulnerable.

03. How can we use data to identify specific days and locations with anomalous weather patterns?

Project Objective

To analyse over a decade of Sri Lankan weather data to detect anomalies using an Isolation Forest model and identify potential indicators of climate change.



Dataset

- **Source:** SriLanka_Weather_Dataset.csv
- **Time Span:** Jan 2010 – Jun 2023 (13.5 years)
- **Scope:** 147,480 daily records across multiple cities.
- **Key Metrics:** Temperature, Precipitation, Wind Speed, Weather Codes.

Methodology

01.
Data Preparation

- Loaded and cleaned the dataset.
- **Result:** No missing values found.
- Converted date columns for time-series analysis.

02.
Exploratory Analysis

- Analysed distributions and trends.
- Identified key weather patterns and seasonality.

03.
Feature Engineering

- Created time-aware features to improve model accuracy.
- **e.g.,** 7-day rolling temperature average, lagged variables.

04.
Anomaly Detection Modelling

- Trained an Isolation Forest model to identify statistical outliers in weather patterns.

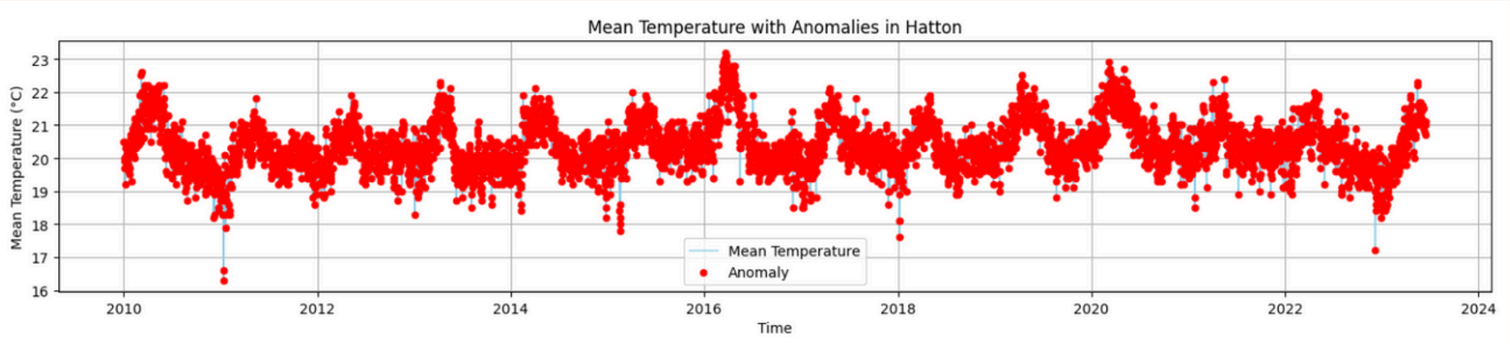
20,065

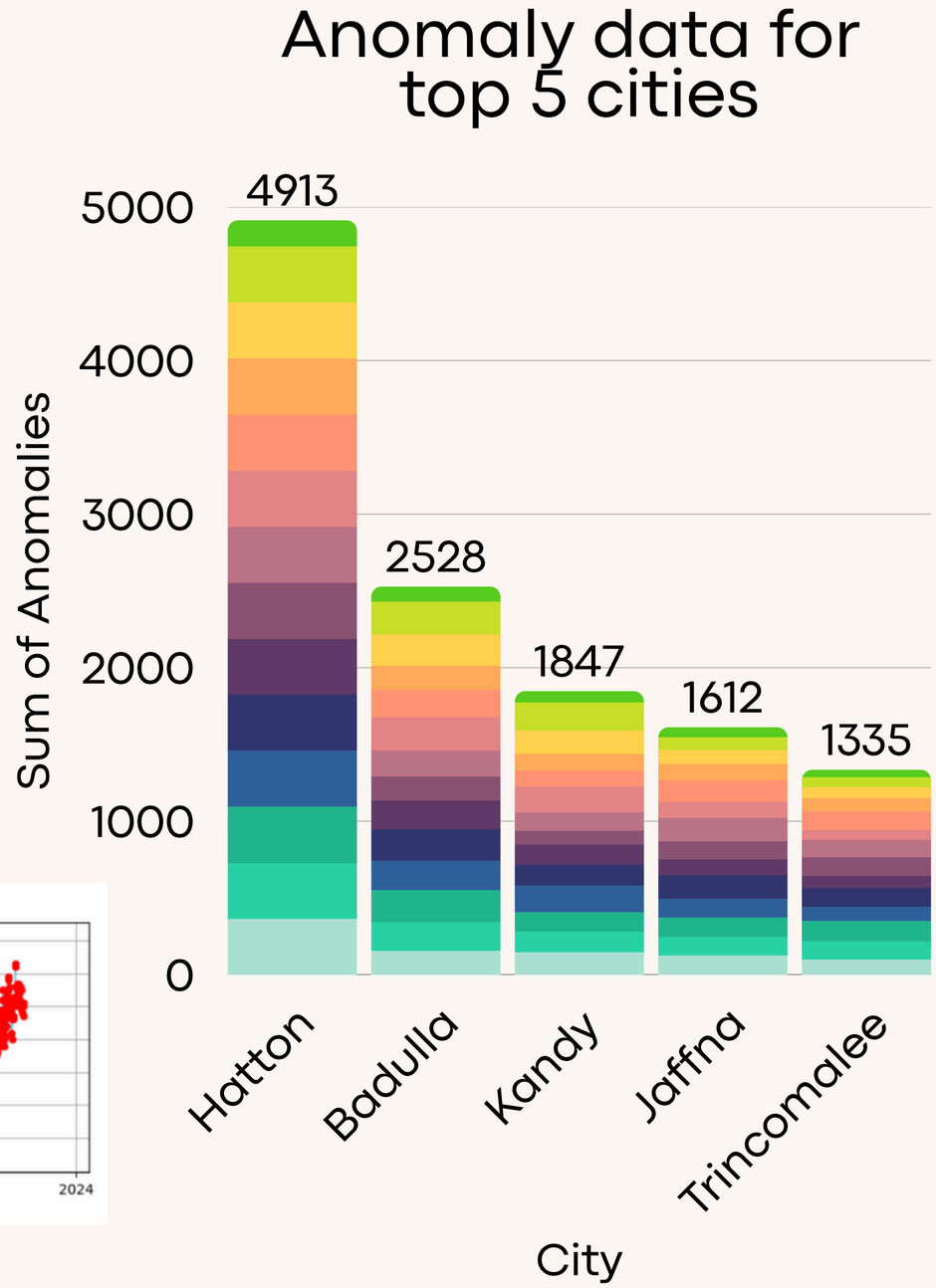
Anomalies Detected

The model flagged ~14% of the days in the dataset as statistically unusual, indicating significant weather volatility.

Geographic Concentration

- Anomalies were not random; they were concentrated in specific cities.
- **Top 5 Cities:** Hatton, Badulla, Kandy, Jaffna, and Trincomalee.

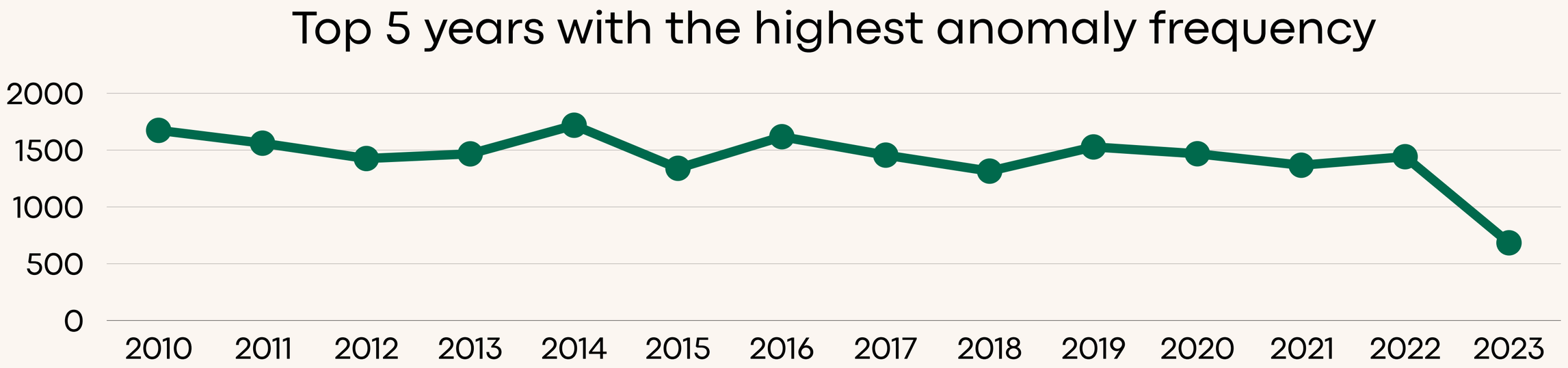




colab.research.google.com

Temporal Patterns

- Certain years showed a higher frequency of anomalies.
- **Top Years:** 2014, 2010, and 2016, 2011, 2019.



colab.research.google.com

Anomaly Profile

Anomalous days are characterized by:

- Extreme precipitation (nearly 3x higher than average).
- Higher wind speeds.
- Greater temperature variance (hotter highs, colder lows).

Conclusion & Next Steps

Conclusion

- The Isolation Forest model successfully identified and located thousands of climate anomalies in Sri Lanka.
- The analysis provides a data-driven map of regional climate instability, highlighting specific cities and years that require further attention.
- These flagged anomalies serve as strong potential indicators of climate change impact.

Next Steps & Future Work

- Analyze Anomaly Severity: Quantify how extreme each anomaly is to classify events by risk level.
- Correlate with Climate Events: Link findings to global phenomena like El Niño cycles to uncover root causes.
- Explore Advanced Models: Use models like STL decomposition to validate and expand upon these initial findings.

Data Mining Individual Project - Final Presentation
Advanced Data Mining - CS5803,
Master of Data Science and AI - CSE, UOM.

Thank you very much!

Deeghayu Adhikari - 258722D
20.07.2025