Influence of Solar Activity on Earth's Climate

Redwanul Karim - 23426184

Research Question

Is there any relationship between solar activity (Solar Flares) and climate change on Earth?

Introduction

Importance of Study

- Understanding the influence of solar activity on Earth's climate is import for predicting weather patterns and long-term climate change.
- Solar activity, including phenomena like solar flares, can impact Earth's atmosphere, potentially affecting temperature and atmospheric composition

Objectives

- Assess whether there is a relationship between solar activity (specifically solar flares) and changes in Earth's climate, focusing on temperature and CO2 concentration.
- Utilize statistical methods to analyze the significance of solar flare data in relation to climate variables.
- Use advanced data engineering methods to Extract, Transform, and Load the data for analysis.

Data Used

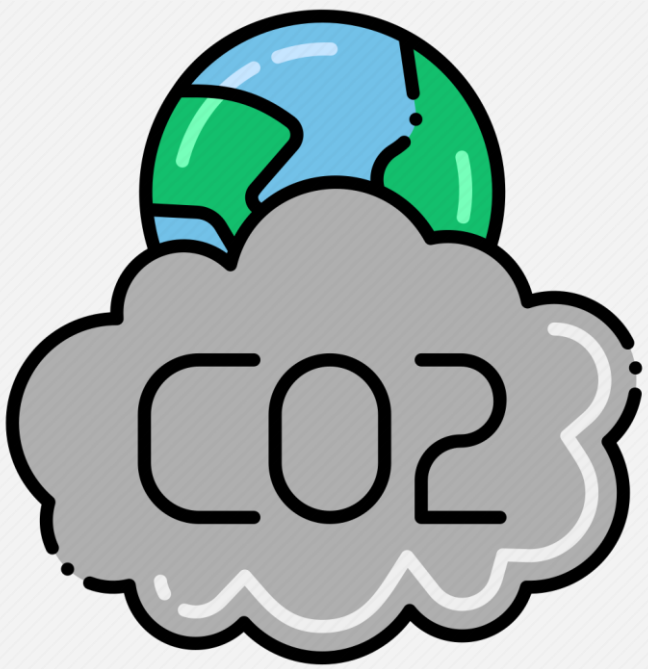
Solar Flare Data

CO2 Concentration Data

Temperature Change Data

Data Used

- **Solar Flare Data**
  - **Source:** Zenodo
  - **Used features**: Date, FlareNumber, TOTUSJH, TOTBSQ, TOTPOT, TOTUSJZ, ABSNJZH, SAVNCPP, USFLUX
  - **Description**: Comprising 8,874 records spanning from May 2010 to December 2019, provides crucial insights derived from vector magnetic field data.
  - **Data Structure and Quality**: Tabular format (CSV), sourced from reliable JSOC and SWPC
  - **Licensing**: MIT License



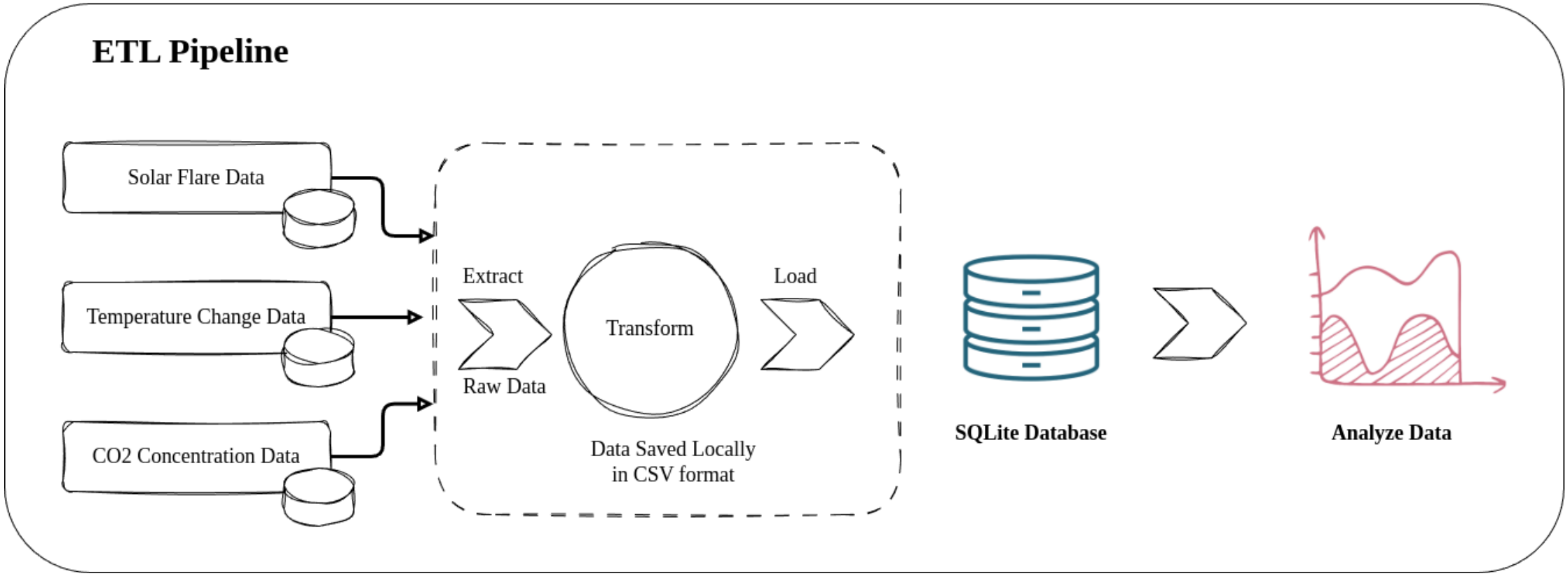
Data Used

- **CO2 Concentration Data**
  - **Source**: Atmospheric CO2 Concentrations from IMF
  - **Used features**: Date, CO2_Concentration_PPM
  - **Description**: The Atmospheric CO2 dataset offers monthly and yearly records of carbon dioxide levels in the air dating back to 1958, enabling users to track changes over time.
  - **Data Structure and Quality**: Tabular format (CSV), sourced from FAOSTAT
  - **Licensing**: IMF License

Data Used

- **Temperature Change Data**
  - **Source:** Annual Surface Temperature Change from IMF
  - **Used features:** Date, Temp_Change
  - **Description:** Shows how Earth's average surface temperature has changed from 1961 to 2021 compared to temperatures between 1951 and 1980, using data from NASA GISS
  - **Data Structure and Quality**: Tabular format (CSV), sourced from reliable JSOC and SWPC
  - **Licensing:** IMF License

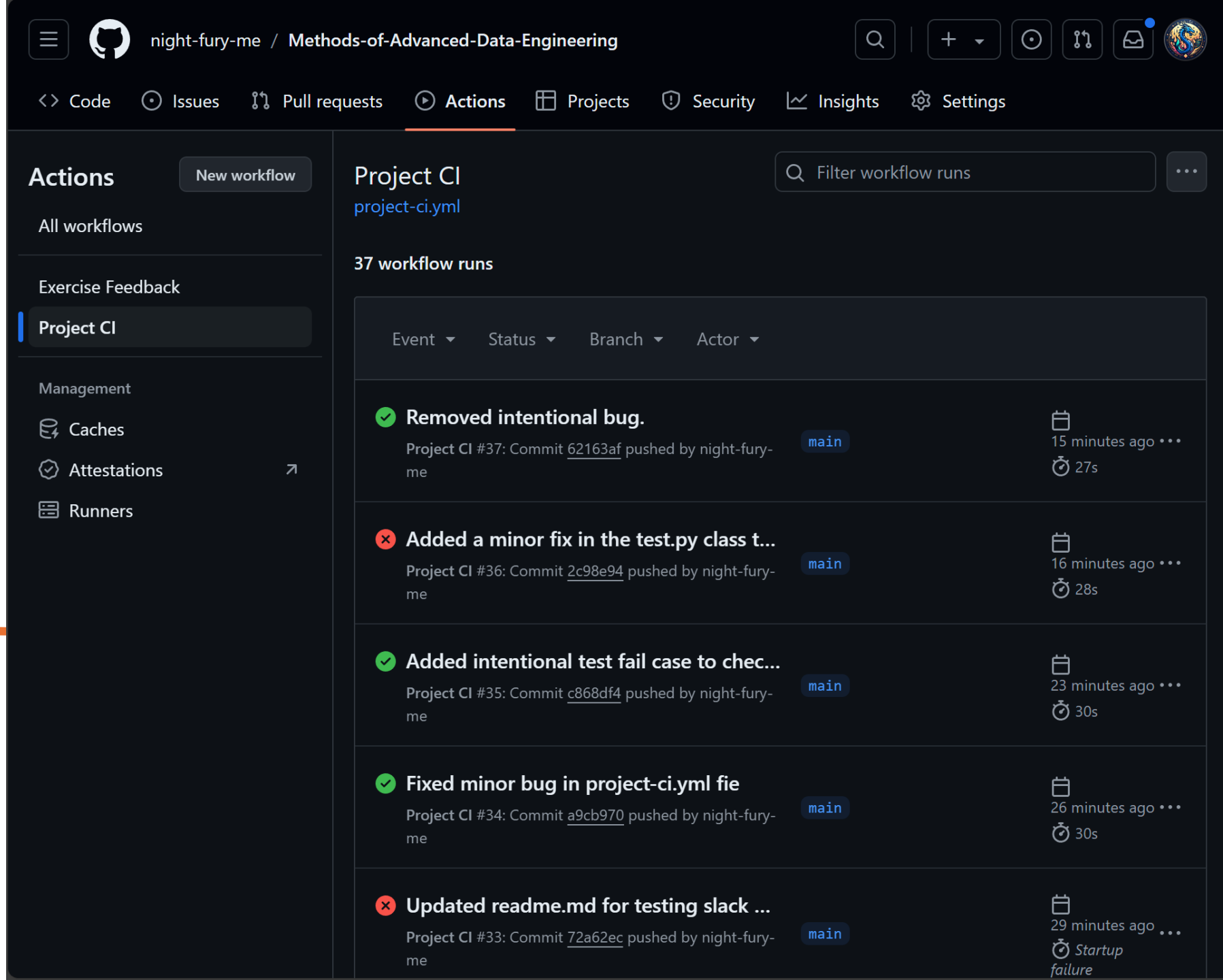
Data Pipeline (ETL)



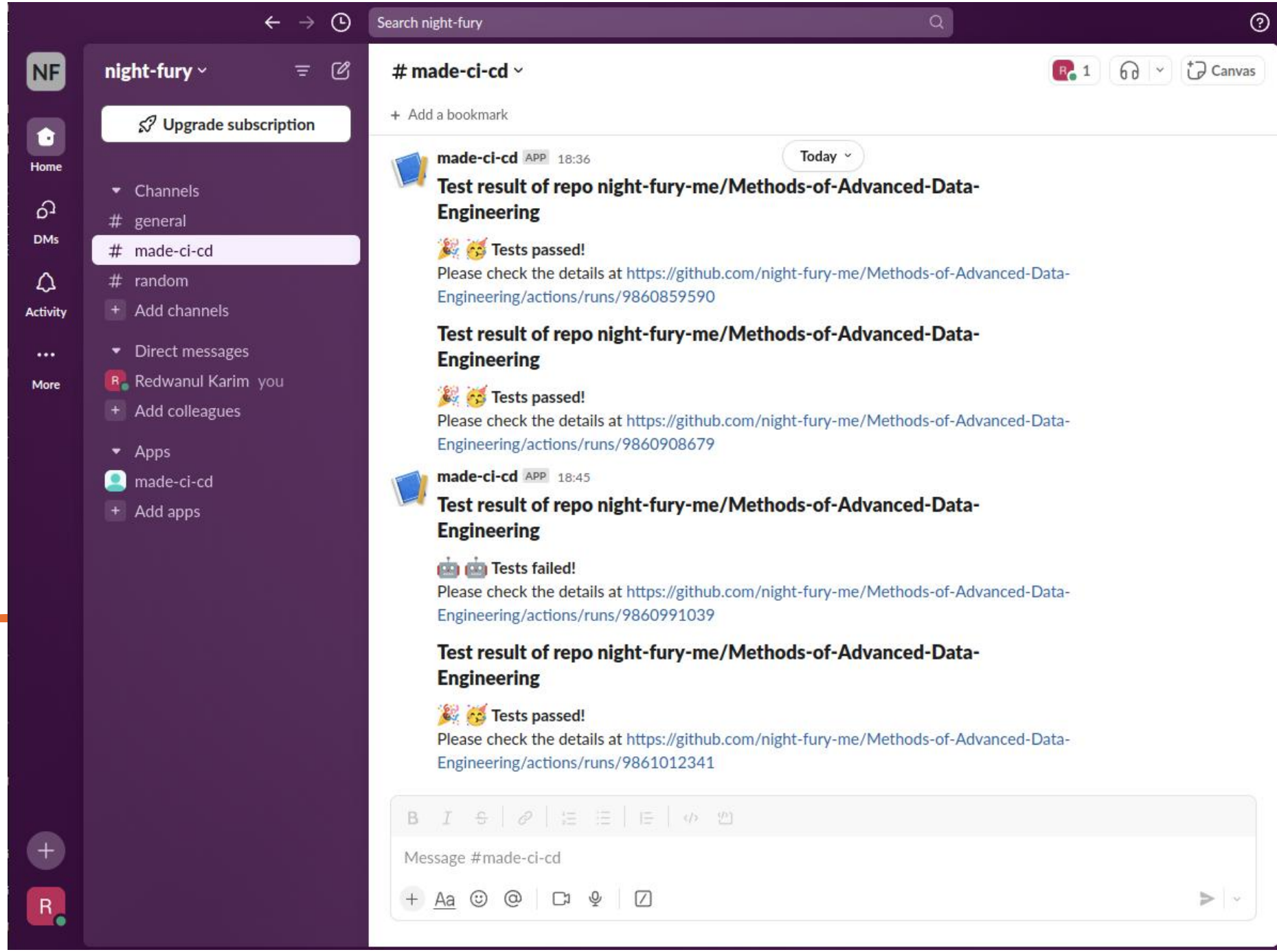
Project Structure

```
project
├── ETL
│   ├── __init__.py
│   ├── extract
│   │   ├── __init__.py
│   │   ├── csv_extractor.py
│   │   ├── extractor.py
│   ├── load
│   │   ├── __init__.py
│   │   ├── loader.py
│   │   └── sqlite_loader.py
│   └── transform
│       ├── __init__.py
│       ├── co2_concentration_data_transformer.py
│       ├── solar_flare_data_transformer.py
│       ├── temperature_data_transformer.py
│       └── transformer.py
├── config
│   └── pipeline_config.yaml
├── data
│   ├── raw
│   ├── sink
│   └── transformed
├── logger
│   ├── __init__.py
│   ├── base_logger.py
│   ├── console_logger.py
│   ├── file_logger.py
│   ├── logger.py
├── logs
│   └── 20240701_043653.log
├── logs
│   └── 20240701_043653.log
├── tasks
│   ├── __init__.py
│   └── task.py
├── tests
│   ├── __init__.py
│   ├── mock
│   │   ├── __init__.py
│   │   ├── data
│   │   │   ├── __init__.py
│   │   │   ├── raw
│   │   │   ├── sink
│   │   │   ├── co2_concentration_mock_data.py
│   │   │   ├── solar_flare_mock_data.py
│   │   │   ├── temperature_change_mock_data.py
│   │   │   └── transformed
│   │   └── mock_logger.py
│   ├── test_system.py
│   └── test_transform.py
├── utils
│   ├── __init__.py
│   ├── config.py
│   └── converters.py
├── pipeline.py
├── tests.py
└── requirements.txt
```

Test Automation with GitHub actions



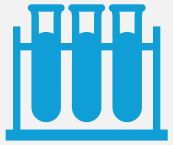
Test Automation with GitHub actions



Methodology

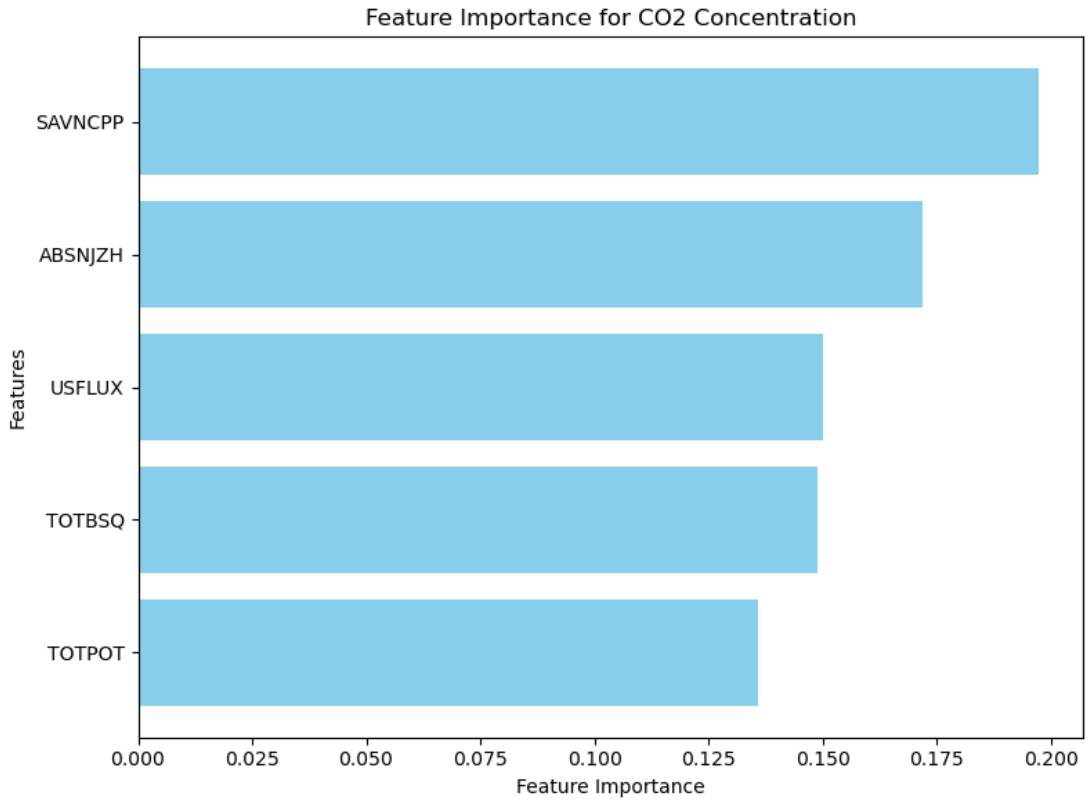
Determining the common time window of the datasets

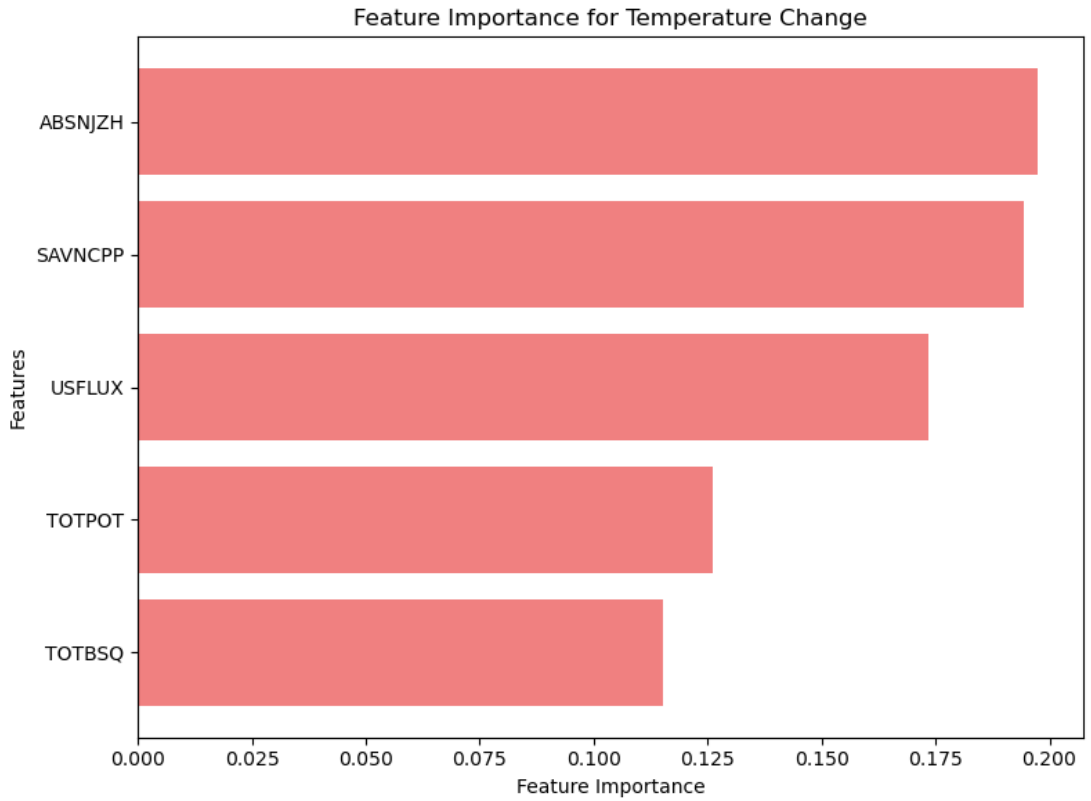
Calculate Feature Importance using Random Forest Regressor



Perform hypothesis testing using t-tests

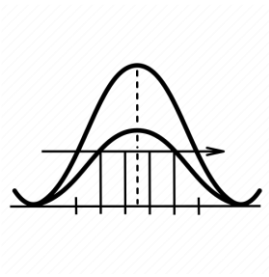
Feature Importance



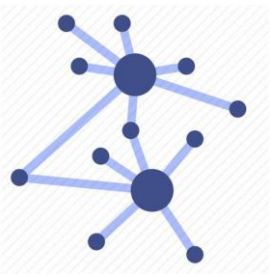


Results

Statistically significant p-values for both CO_2 concentration and temperature change suggest a potential relationship between solar flare events and climate variables on Earth.

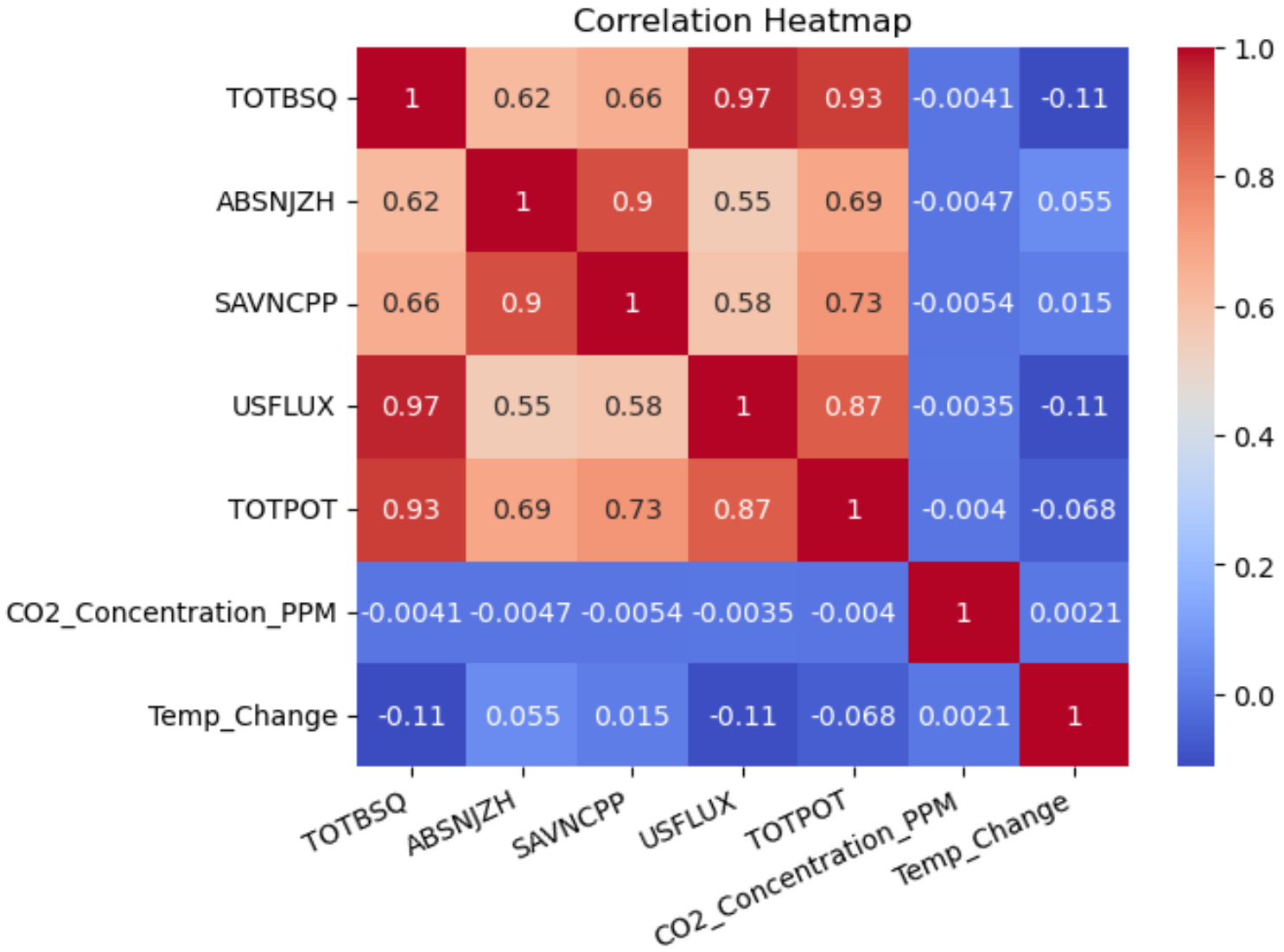


However, none of the datasets are normally distributed, indicating violations of the underlying assumptions of normality.

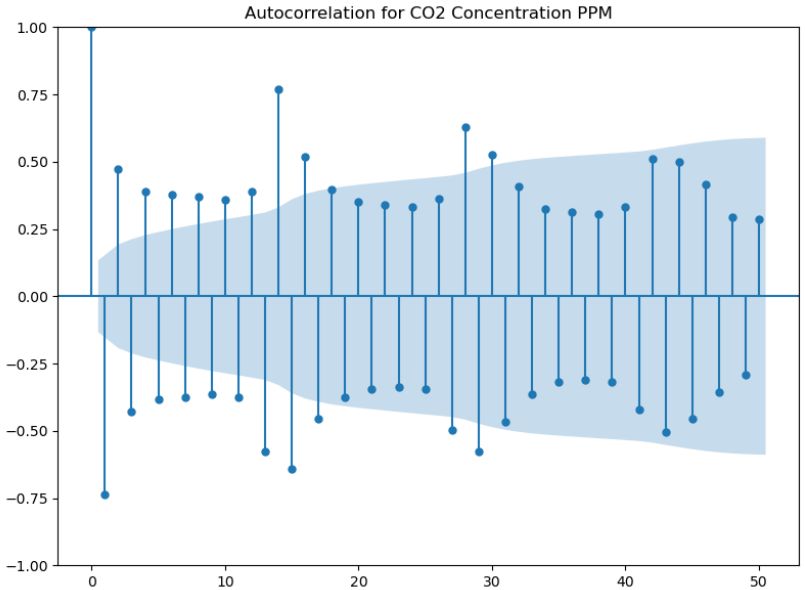


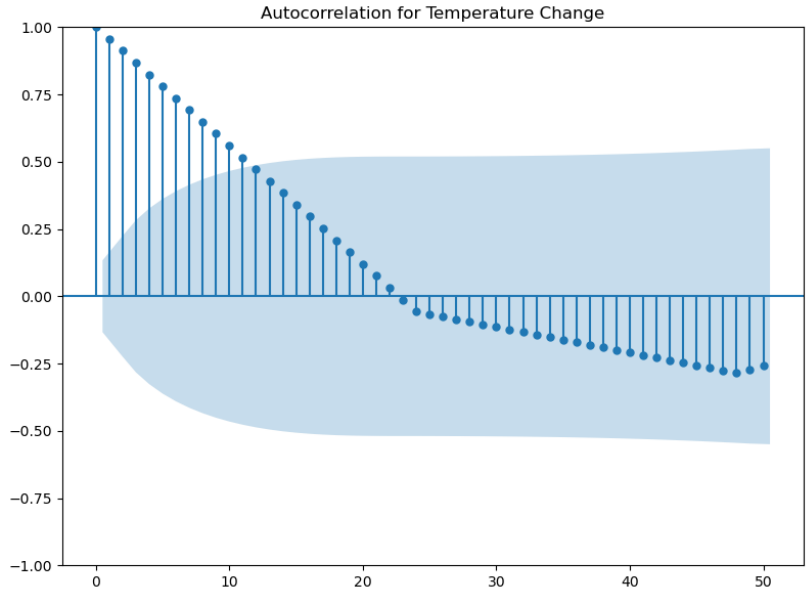
Also, the issue of autocorrelation among data points violates the assumption of homogeneity of variance, raising concerns about the reliability of our hypothesis test results.

Results: Correlation Coefficient Heatmap



Results: Auto correlations

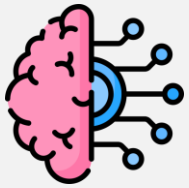




Limitations and Future Direction

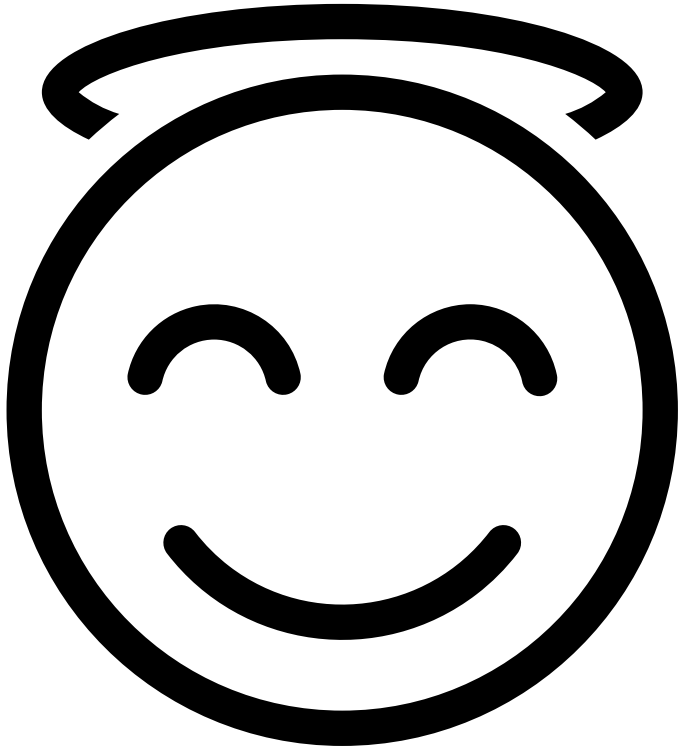
This study relies on publicly available datasets, which may have inherent biases or limitations in data quality and completeness.

The use of traditional statistical tests assumes linear relationships and strict adherence to assumptions like normality, which may not fully capture complex interactions in climate and solar data.



Employ advanced machine learning techniques that can handle non-linear relationships and complex data interactions to refine our understanding of the relationship between solar activity and climate change.

Thank you for your attention!



Reference

- Solar Flare Data by Zenodo https://zenodo.org/records/4603412
- Temperature Change Data by IMF https://climatedata.imf.org/datasets/4063314923d74187be9596f10d034914/explore
- CO2 Concentration Data by IMF https://climatedata.imf.org/datasets/9c3764c0efcc4c71934ab3988f219e0e/explore