

# Methods of Advanced Data Engineering

## "Climate Change: Greenhouse Gas(Co<sub>2</sub>) Emissions and Temperature Trends in Rathaus area of Konstanz"

Ferdoushi Nahar(23129118)

# Content

- 01 Introduction
- 02 Project Structure
- 03 Data Sources
- 04 Data Pipeline
- 05 Analysis
- 06 Conclusions
- 07 References

# 1. Introduction

**Objective:** Investigate the relationship between CO<sub>2</sub> emissions and temperature trends in the Rathaus area of Konstanz.

**Key Question:** How do greenhouse gases (CO<sub>2</sub>) affect climate change in the Rathaus area of Konstanz?

## 2. Structure

The project uses an ETL (Extract, Transform, Load) pipeline structure by following the steps:

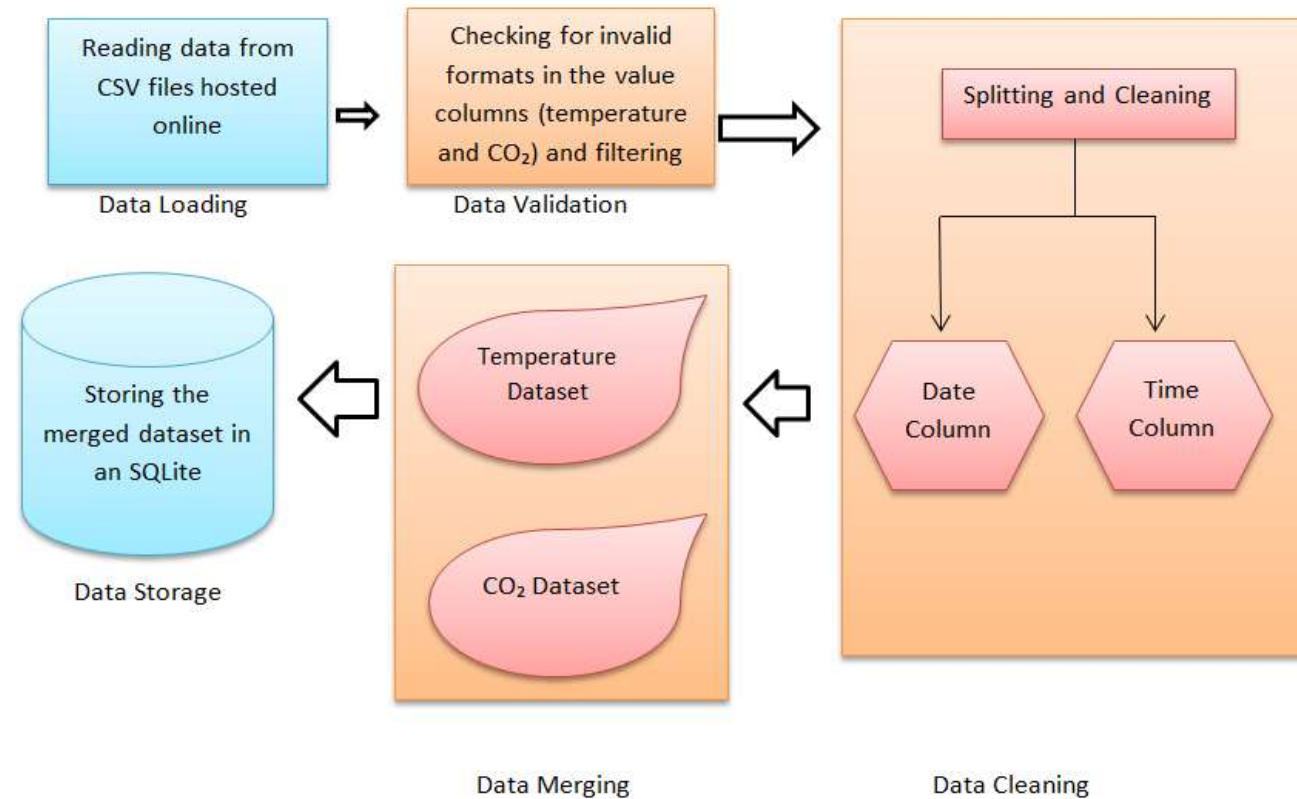
**Extract:** Read data from online CSV files.

**Transform:** Validate: Check for invalid temperature and CO<sub>2</sub> formats.

**Clean:** Split and clean date and time columns.

**Merge:** Combine temperature and CO<sub>2</sub> datasets.

**Load:** Store the merged dataset in an SQLite database



Structure of the project

### 3. Data Sources

**Source of Data:** Open source data

#### **License :**

The datasets are available under an open-data license, which allows for use, distribution, and modification.



Data Source: <https://offenedaten-konstanz.de/>

## Data Source

### **Datasource1: Greenhouse Data**

Data Type: CSV

The dataset contains information on greenhouse gas emissions specifically Co2 in Rathaus area of Konstanz. This data helps in analysing trends in greenhouse gas emissions across Rathaus area of Konstanz over time

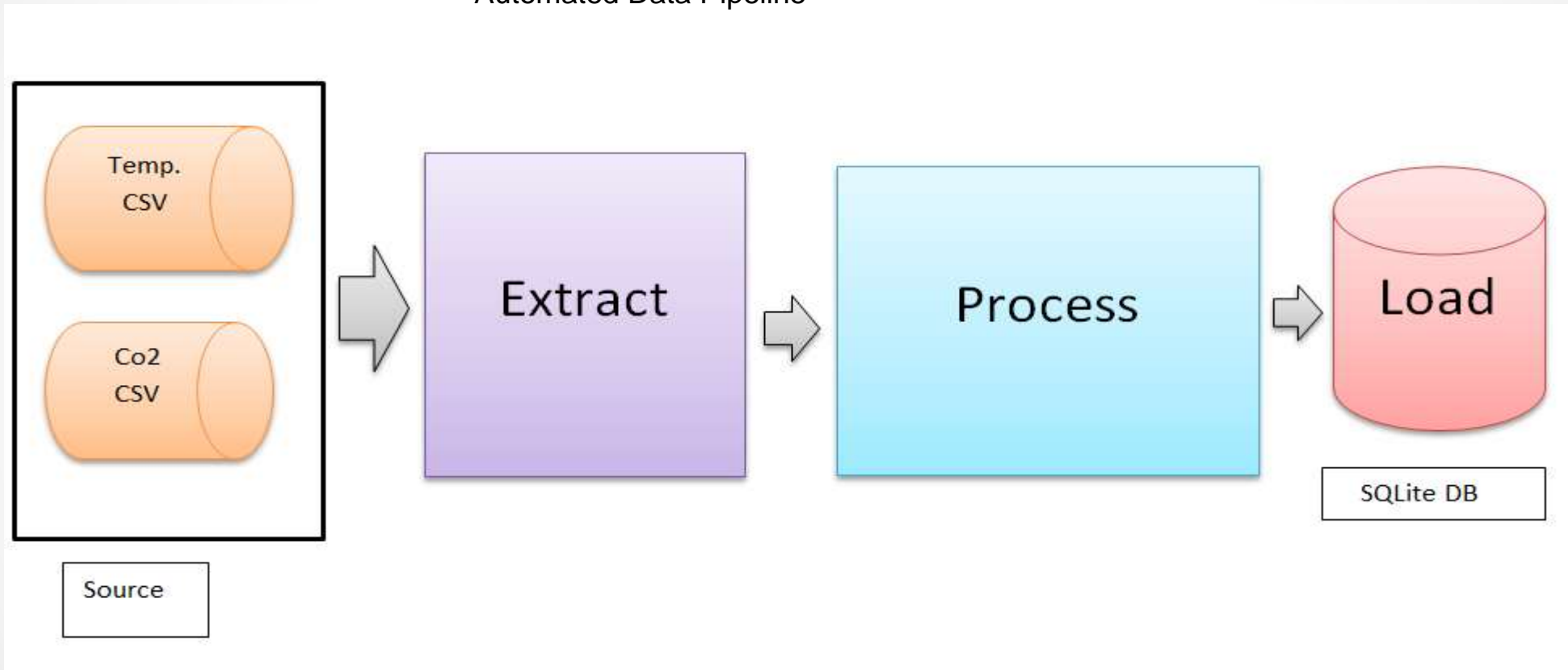
### **Datasource2: temperature data**

Data Type: CSV

The dataset contains information on temperature trend in Rathaus area of Konstanz.

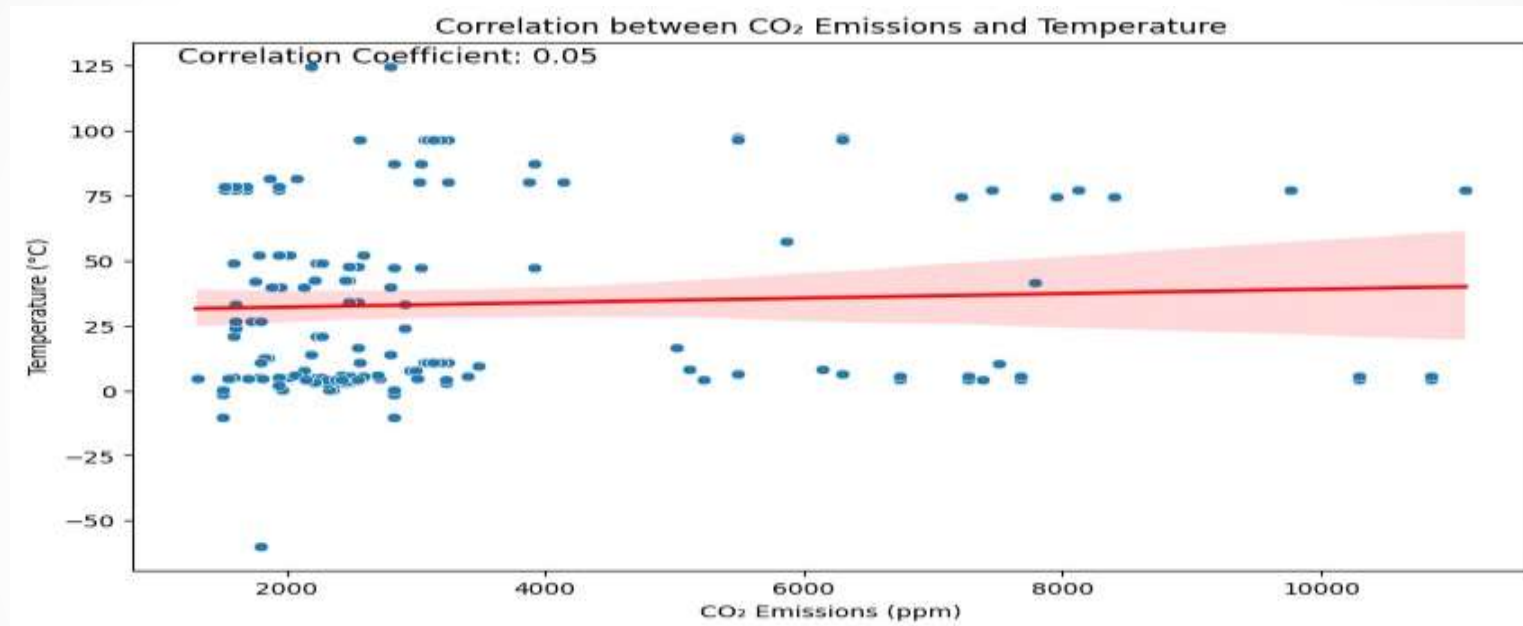
## 4. Data Pipeline

Automated Data Pipeline



## 5. Analysis

### Correlation Analysis:

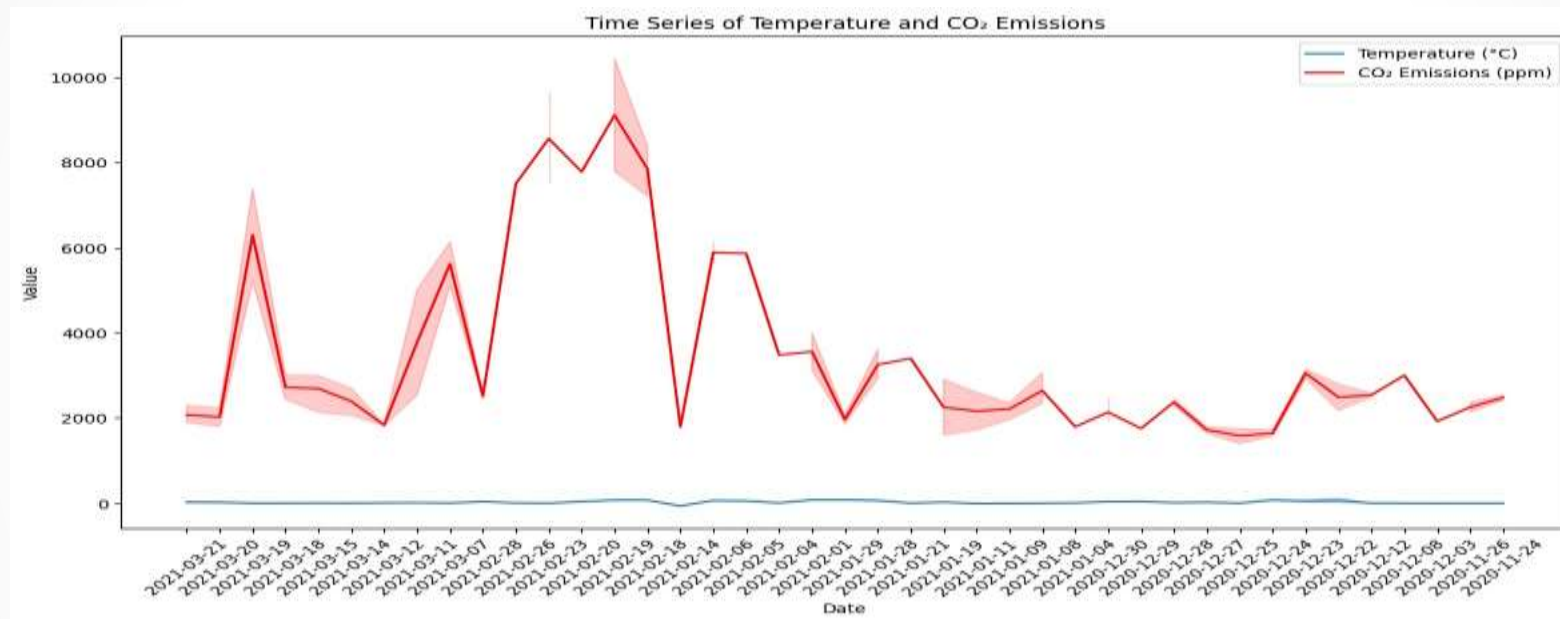


Finding: Weak positive correlation (correlation coefficient: 0.05)



# Analysis

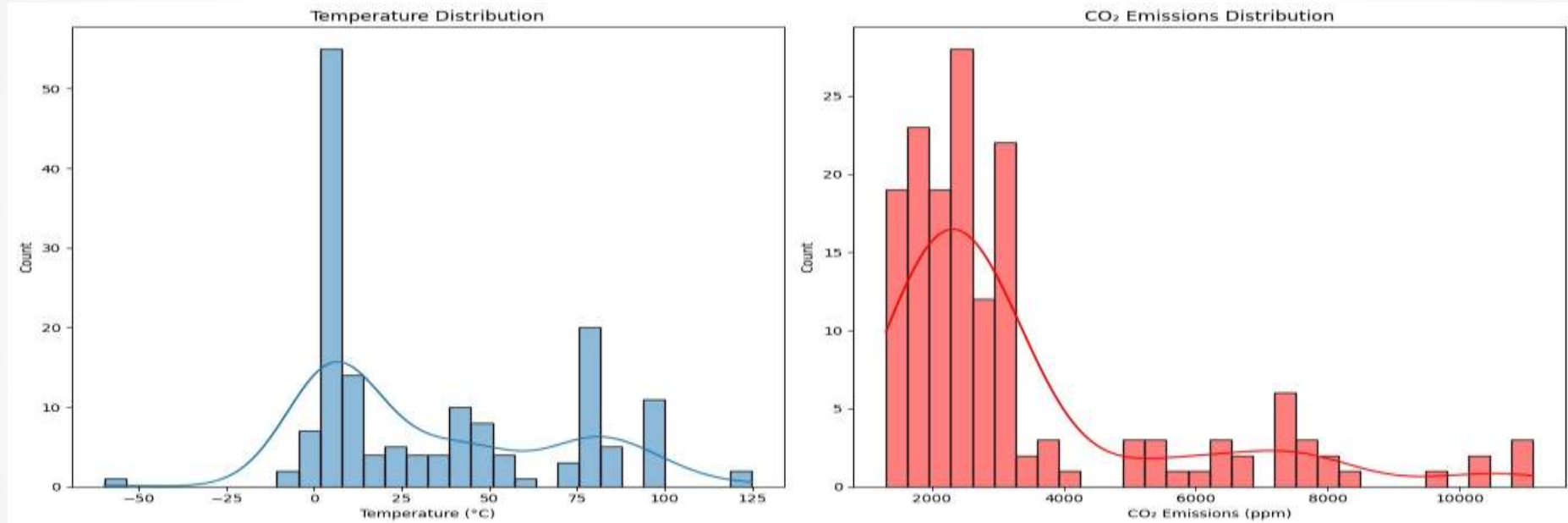
## Time Series Analysis:



- Observation: Temperature remains stable despite CO<sub>2</sub> fluctuations

# Analysis

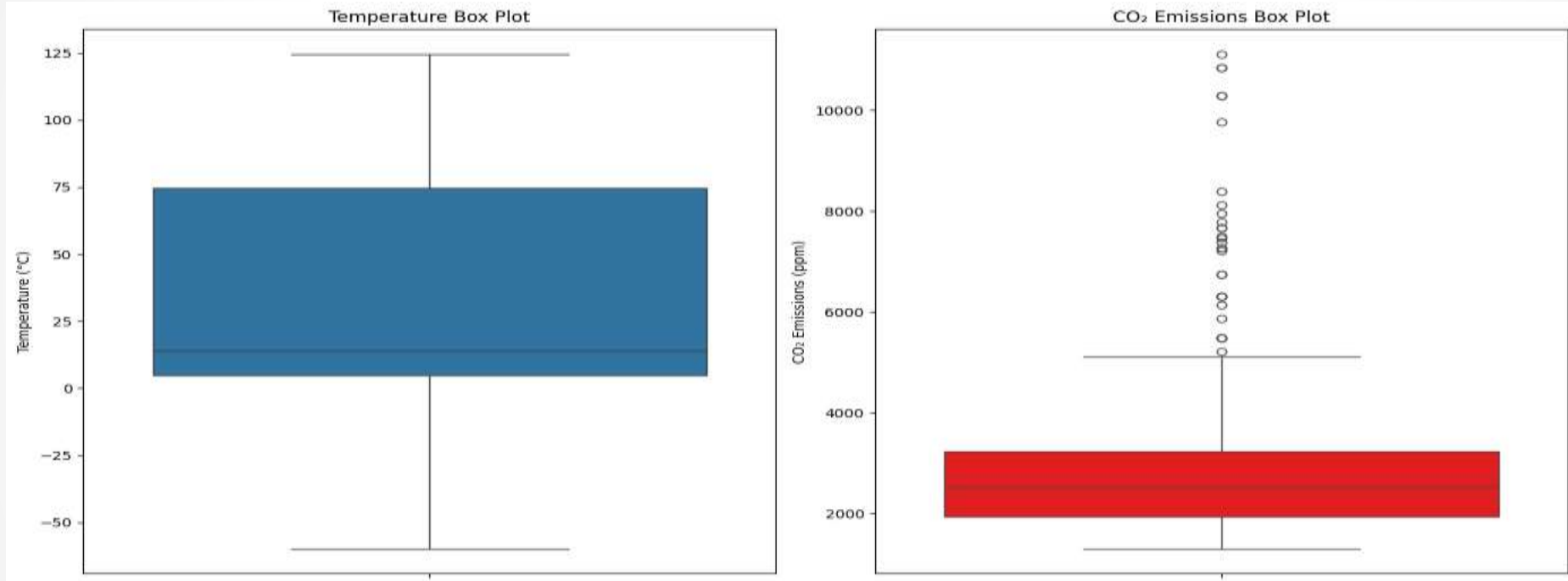
## Distribution Analysis:



- Temperature: Spread out with a lower range peak
- CO<sub>2</sub> Emissions: Highly skewed with values concentrated at the lower end

# Analysis

## Box Plot Analysis:



Observation: CO<sub>2</sub> emissions have many outliers; temperature data is more consistent

## 6. Conclusions

**Conclusion:** No significant linear correlation between CO<sub>2</sub> emissions and temperature trends.

**Limitations:** Does not account for other greenhouse gases or climatic factors

**Future Work:** In Future we need to include other greenhouse gases to check the temperature trends

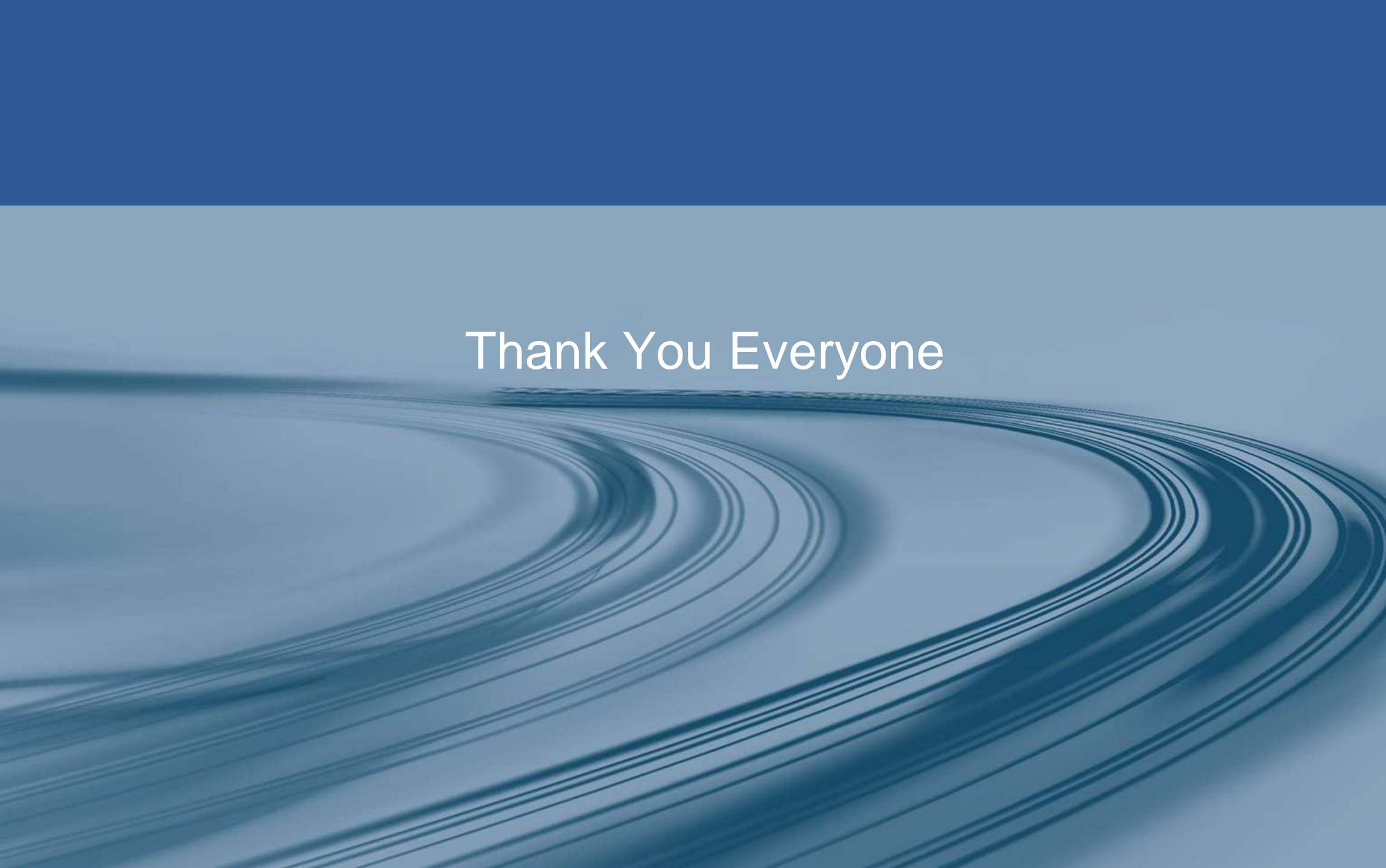
## 7. References

Dataset References:

Greenhouse Data <https://offenedaten-konstanz.de/dataset/co-werte-konstanz>

Temperature Data <https://offenedaten-konstanz.de/dataset/temperaturwerte-konstanz>

Thank You Everyone

The background of the slide is a solid blue color at the top. Below this, there is a large area with a light blue background and a series of concentric, curved lines that create a sense of motion or a ripple effect. The lines are darker blue and curve from the bottom left towards the right side of the slide.

