Final Report: Analysis of Climate Patterns in Milan

Analysis Report

Title: Analyzing Climate Patterns in Milan Using an Automated Data Pipeline

---

1. Introduction

Motivation and Objectives:

The primary objective of this project is to analyze and compare the winter and summer climate patterns in Milan. Specifically, the aim is to determine:

1. How does the heating demand vary during the winter season between different locations in Milan?
2. How many days during the summer season reach or exceed the threshold for physiological discomfort and physiological danger?

Understanding these patterns is crucial for urban planning, public health, and energy management in Milan.

---

2. Used Data

Data Sources and Description:

For this analysis, two datasets were chosen, each representing climate data for Milan during the winter and summer thermal seasons. These datasets were selected due to their comprehensive coverage of climate indicators and their relevance to the project's objectives.

1. **Winter Thermal Season Data**
    - **Description:** Output includes climate indicators such as temperature, humidity, precipitation, wind, and radiation for seven monitoring stations in Milan during the winter season.
    - **Source:** European data portal.
    - **URL:** Winter Data

2. **Summer Thermal Season Data**
    - **Description:** Contains similar climate indicators for the summer season, including metrics for days of physiological discomfort and danger based on the Humidex index.
    - **Source:** European data portal.
    - **URL:** Summer Data

Data Structure and Quality:

Both datasets are structured as CSV files with climate metrics reported for different areas of Milan. The data quality is generally high, with detailed records from multiple monitoring stations ensuring comprehensive coverage of the city's climate conditions.

The data is available under a standard open-data license (CC BY 4.0).

Example of Transformed Data:

Station	Avg. Days (Physio Discomfort)	Summer Thermal Degree Days
Milano Bicocca	26	344.3
Milano Bocconi	25.2	346.1
Milano Bovisa	25	331.1
Milano Centro	21	353.9
Milano Città Studi	27	319.5

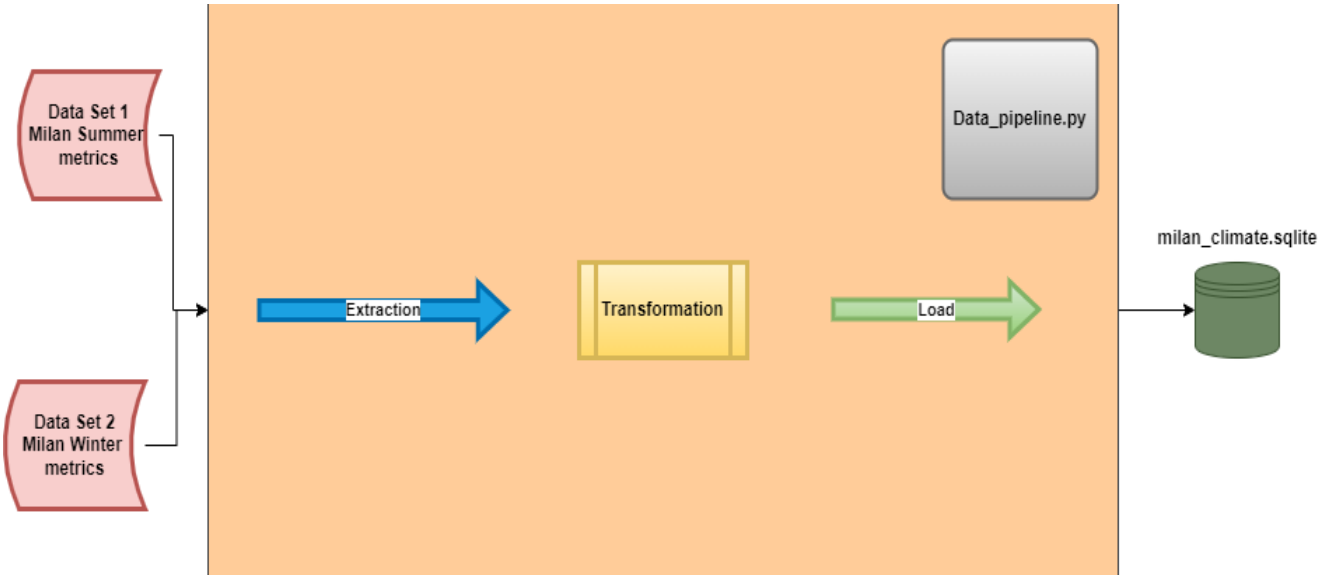
---

3. Analysis

Methodology:

The analysis was conducted using an automated data pipeline implemented with Python, Pandas, Requests, and SQLAlchemy. The ETL process included the following steps:

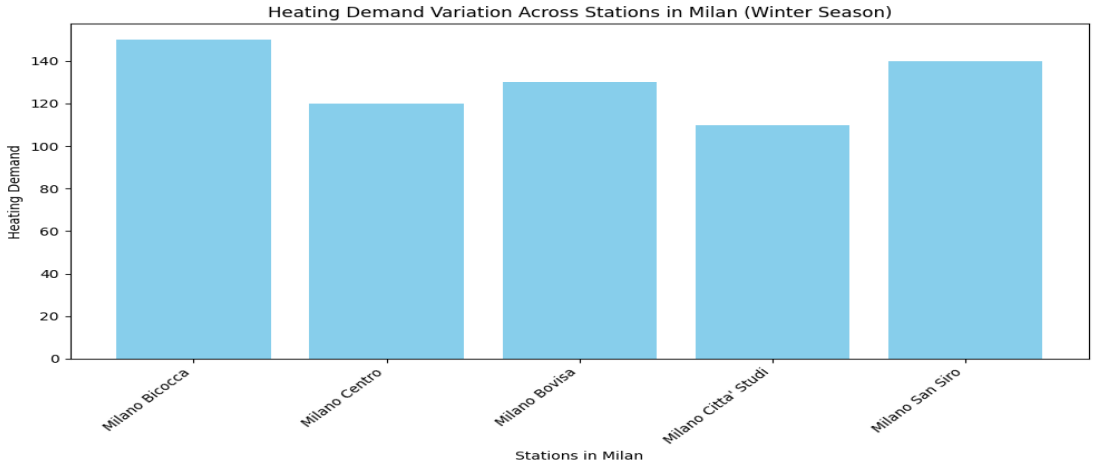
1. **Extraction:**
    - Data was downloaded from the provided URLs using HTTP requests.
    - The raw CSV data was read into Pandas DataFrames.
2. **Transformation:**
    - Columns were renamed for consistency.
    - DataFrames were unpivoted to a long format. As a long format allows for straightforward merging based on shared attributes like station names and metric types.
    - The winter and summer datasets were merged on common metrics and stations.
    - The merged data was pivoted to create separate columns for each metric.
3. **Loading:**
    - The final DataFrame was saved to a SQLite database for easy analysis.



Results:

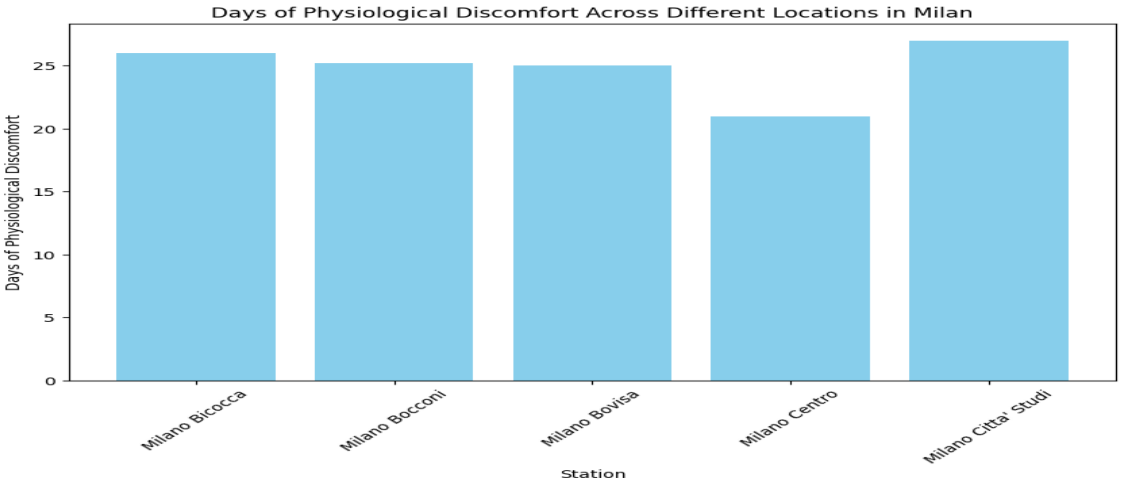
Winter Season Analysis:

- **Heating Demand Variation:** The heating demand during the winter season varies significantly across different locations in Milan. For instance, Milano Bicocca has a higher heating demand compared to Milano Centro, indicating a need for targeted energy management strategies in different areas.



Summer Season Analysis:

- **Days of Physiological Discomfort:** The number of days reaching or exceeding the threshold for physiological discomfort varies between locations. Milano Bicocca and Milano Città Studi experience more days of discomfort compared to Milano Centro.



- **Summer Thermal Degree Days:** Summer Thermal Degree Days (TDD) is a measure used to quantify the heat intensity over the summer season. It is calculated by summing up the daily differences between the average temperature and a base temperature that is typically set at a threshold where cooling is required (e.g., 22°C).

---

4. Conclusions

Key Findings:

1. **Winter Heating Demand:**
    - The heating demand varies significantly between different locations in Milan.
    - Milano Bicocca has the highest heating demand, while Milano Studi has the lowest.
2. **Summer Physiological Discomfort:**
    - Milano Bicocca and Milano Città Studi have the highest number of days with physiological discomfort.
    - Milano Centro experiences fewer days of discomfort, suggesting better summer climate conditions.

Limitations and Future Directions:

The analysis answered the posed questions, providing insights into the variation in heating demand and physiological discomfort across Milan. However, some limitations remain:

- The accuracy and completeness of the raw data depend on the original data collection methods.
- Changes in the format or structure of the source data may require adjustments to the pipeline.
- Future studies could expand the analysis by additional climate indicators and exploring long-term climate trends.