**tENERGY CONSUMPTION OPTIMIZATION USING IOT AND BIG DATA**

**OMARI GALANA SHEVO-CT204/108375/21**

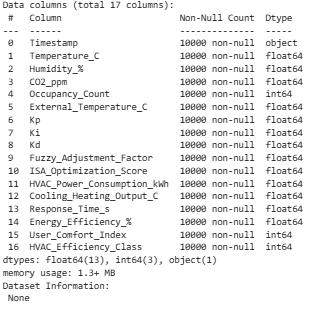
**HVAC Dataset Analysis Report**

# 1. Data Overview

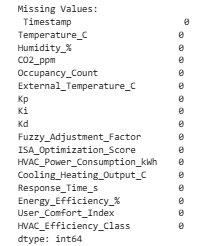
The analysis begins with an exploration of the HVAC dataset, which includes various features related to energy consumption, environmental conditions, and system performance. The goal is to understand the relationships between energy consumption and other variables, including the PID parameters used in the HVAC system.

## **1.1 Data Loading and Initial Inspection**

The dataset was successfully loaded using pandas and displayed the first few rows (head) and last few rows (tail) to inspect its content. We also examined the data types and checked for any missing values in the dataset.



Missing Values: A quick check confirmed no missing data, ensuring the dataset is complete for analysis



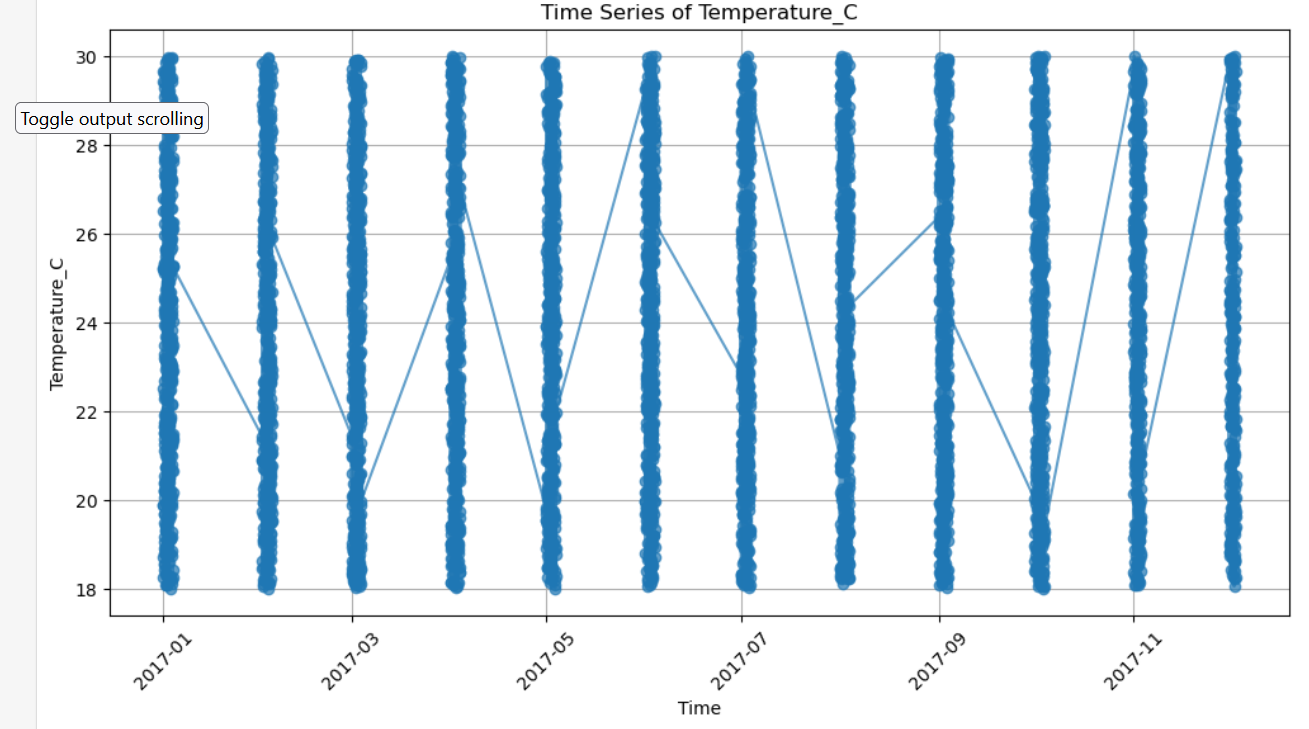
## **1.2 Key Features and Units**

The following key features were identified:

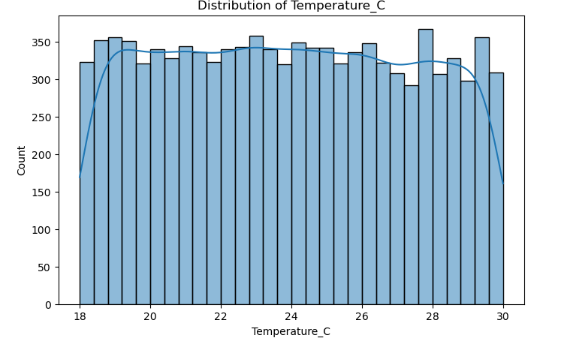
* Energy Consumption: Measured in kWh ('HVAC\_Power\_Consumption\_kWh') and efficiency percentage ('Energy\_Efficiency\_%')
* Environmental Factors: Temperature (°C), Humidity (%), CO2 (ppm), Occupancy Count, External Temperature (°C)
* System Performance Features: PID controller parameters (Kp, Ki, Kd), fuzzy adjustment factors, response times (in seconds), and system performance scores

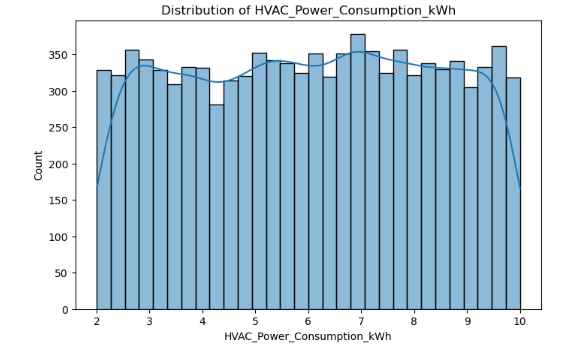
2. **Data Visualization**

2.1 Time Series Analysis



We visualized the time series of key features, including HVAC power consumption, internal and external temperatures, and humidity levels: Here are some of the visualizations:



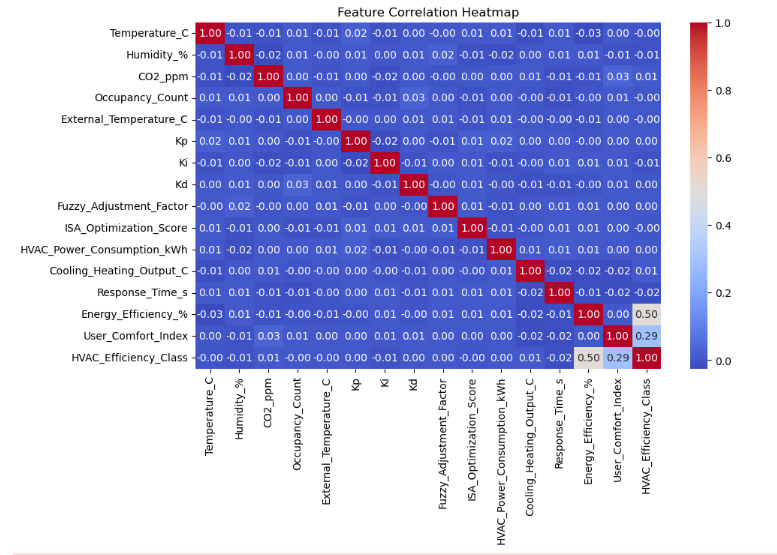


# 3. Data Analysis

## **3.1 Correlation Analysis**

Pearson correlation coefficients were calculated to examine the relationship between HVAC power consumption and other numerical variables. The results indicated weak or no strong correlations between HVAC energy consumption and most other features.

The graph below is a feature correlation heatmap from the dataset.



Key Observations:

No strong correlation between energy consumption and environmental variables such as temperature and humidity.

System performance features (PID parameters) also showed weak correlation with energy consumption.

## **3.2 High Energy Consumption Analysis**

By focusing on the top 5% of energy consumption instances, we observed subtle differences in several variables. However, no significant patterns emerged to explain high energy consumption beyond what was observed in the overall dataset.

## **3.3 PID Parameter Impact**

The impact of PID controller parameters (Kp, Ki, Kd) on energy consumption was analyzed by grouping the parameters into quantiles. This analysis did not reveal any clear trends linking PID adjustments to energy usage. The relationship remains unclear, and further investigation is needed to identify potential impacts.

**4. Challenges encountered:**

**Data complexity**

**Understanding the dataset**

# 5. Summary of Findings

## **5.1 Key Insights**

No strong correlations: There is no strong linear relationship between HVAC power consumption and environmental or system performance features.

High energy consumption periods: The top 5% of high energy consumption periods did not show distinct patterns that could be attributed to specific variables.

PID parameters and energy consumption: The analysis of PID parameters (Kp, Ki, Kd) using quantile-based binning did not provide conclusive evidence of their impact on HVAC energy consumption.

## **5.2 Next Steps**

Non-linear Relationships: The analysis primarily focused on linear correlations. It would be beneficial to explore non-linear relationships using more advanced techniques, such as scatter plot matrices or non-linear correlation methods.

Feature Engineering: Additional feature engineering and the application of machine learning models (such as regression) could be considered to better understand the interactions between variables and improve energy consumption predictions.

Further Investigation of PID Parameters: A more granular analysis of the PID parameters, perhaps using different control system analysis techniques, could yield better insights into their effects on energy consumption.

# 6. Conclusion

In conclusion, while basic statistical analysis and visualizations were insightful, they did not reveal strong, actionable insights regarding HVAC energy consumption. The weak correlations between energy consumption and other variables suggest the need for more advanced techniques, such as machine learning, to uncover the complex relationships in the data. Further analysis of PID parameters, considering different control techniques, may also provide more clarity regarding their impact on energy usage.

**7. Timeline**

|  |  |
| --- | --- |
| March 10th | Finalize building the model |
| March 12th | Deploying the model |