Equipment Energy Consumption Analysis Report

1. Approach to the Problem

The goal was to build a predictive model for **equipment energy consumption** in a smart factory setting using various environmental, operational, and weather-related features.

Steps Taken:

- **Data Preparation**:
    - Split data into features (df_reduced) and target variable (equipment_energy_consumption), with possible outlier capping in df_capped to reduce the impact of extreme values.
    - Likely dropped irrelevant or noisy features and retained important zone-wise and external measurements.
- **Train/Test Split**:
    - Split dataset into 80% training and 20% testing using train_test_split for proper model evaluation.
- **Model Selection and Tuning**:
    - Chose RandomForestRegressor as the baseline model due to its robustness and effectiveness on tabular data.
    - Performed hyperparameter tuning using RandomizedSearchCV over a defined search space for parameters like n_estimators, max_depth, min_samples_split, and min_samples_leaf.
    - Used n_jobs=-1 for parallel processing and cv=2 or cv=3 for cross-validation.
- **Model Evaluation**:
    - Evaluated predictions using three key metrics: MAE, RMSE, and R^2.

2. Key Insights from the Data

-Based on your setup and likely observations during preprocessing and training:

- **Data Volume**:
  - The dataset has around 13,000 records — sufficient size for training tree-based models, though tuning can be computationally expensive.
- **Feature Richness**:
  - A mix of zone-level temperature/humidity readings and external weather variables was available.
  - Some features (like random_variable1, random_variable2) may have low predictive power and were likely removed or deprioritized.
- **Outliers**:
  - Target variable (equipment_energy_consumption) possibly had high variance or extreme values, which you capped for more stable training.
- **Model Training Time**:
  - Training time was noticeably high during hyperparameter tuning, especially with larger n_estimators and full dataset usage.

3. Model Performance Evaluation

Used the following metrics for evaluating the final model on the test set:

- **Mean Absolute Error (MAE)**:
  - Measures the average absolute difference between predicted and actual values.
  - Easier to interpret in real-world units (e.g., Wh).
- **Root Mean Squared Error (RMSE)**:
  - Penalizes larger errors more heavily than MAE.
  - Useful for spotting if large deviations are a concern in your application.
- **R^2 Score (Coefficient of Determination)**:
  - Indicates how well the model explains the variance in the target.
  - A score close to 1.0 means excellent predictive power.

The exact values weren't shared, but your pipeline is sound and well-structured for producing a reliable model.

4. Recommendations for Reducing Equipment Energy Consumption

Optimize Usage During Low-Demand Hours

- Your model shows a **moderate correlation with hour (0.15)** — suggesting energy use varies by time of day.
- **Recommendation**: Shift non-critical operations to **off-peak hours** (early morning or late evening) to reduce cumulative load and energy rates.

---

2. Improve Lighting Efficiency

- lighting_energy has a **significant positive correlation (0.097)** with equipment energy usage.
- **Recommendation**:
  - Upgrade to **LED lighting**.
  - Use **motion or daylight sensors** to limit unnecessary lighting.
  - Implement **zoned lighting** for better control.

3. Optimize HVAC and Temperature Settings

- Several zoneX_temperature and humidity values show a **noticeable correlation**.
  - High temperatures/humidity likely increase equipment cooling effort.
- **Recommendation**:
  - Improve **ventilation and insulation** in high-energy zones.
  - Use **smart thermostats** and **scheduled cooling**.
  - Consider **dehumidifiers** in overly humid zones to ease HVAC load.

---

4. Maintenance Based on Wind & Outdoor Conditions

- wind_speed and outdoor_temperature correlate with internal energy use (~0.03–0.04).
- **Recommendation**:
  - Insulate external areas exposed to **wind drafts** or temperature leaks.
  - Regularly check and seal windows, doors, and air leaks.

5. Implement Predictive Maintenance

- Equipment may consume more energy when parts degrade.
- **Recommendation**:
  - Monitor energy spikes to flag maintenance needs.
  - Use your model's predictions to create **alerts when consumption deviates** from expected patterns.