# Problem Statement

**Goal: Predict and optimize zone-level electricity consumption (e.g., for AC units) to reduce energy usage without compromising indoor air quality.**

## CRISP-DM

**CRISP-DM (Cross Industry Standard Process for Data Mining)** is a widely used methodology for structuring data science and machine learning projects. It provides a systematic approach to planning and executing data mining (or data analytics) projects, ensuring that the process is repeatable and understandable.

CRISP-DM consists of **6 phases**:

1. **Business Understanding**
2. **Data Understanding**
3. **Data Preparation**
4. **Modeling**
5. **Evaluation**
6. **Deployment**

## **Phase 1:** **Business Understanding**

### Step 1: Identify the Business Objective

***Business Objective: Predict and optimize electricity consumption (especially AC units) in an office building to reduce energy costs and improve sustainability.***

### Step 2 :Determine Project Goals

| **Business Goal** | **Data Science Goal** |
| --- | --- |
| Reduce AC usage cost | Predict AC consumption in next hour/day |
| Spot unexpected high usage | Detect anomalies in electricity usage |
| Optimize comfort and energy | Build relationship between temperature and AC usage |

➡️ Let's choose a data science goal for now:

❗**Data Science Goal:** Forecast AC unit power consumption in each zone for the next hour using historical sensor and power data.

## **Phase 2: Data Understanding**

We have 14 CSV files (7 floors × 2 years).

The dataset includes:

* Power consumption of AC, lighting, plug loads.
* Temperature, humidity, light (lux).
* Data collected every minute for 18 months.
* Some floors/zones don't have all types of data