# House Price Prediction Project

## Introduction

You are provided with a dataset containing house characteristics and their corresponding prices. Your task is to clean the data and build a model to predict house prices based on the available features. This exercise will help you understand how to deal with real-world data and prepare it for machine learning.

## Objectives

- Identify and fix inconsistencies or missing values in the dataset.  
- Standardize column names and values.  
- Prepare the dataset for modeling.  
- Use machine learning to predict house prices.

## Part 1: Data Cleaning

1. Load and examine the dataset to understand its structure.

2. Clean the column names by removing unnecessary characters or spaces, and make them consistent.

3. Check for missing values in the dataset and decide whether to fill or remove them.

4. Identify inconsistent values in categorical columns (like Location or Furnishing) and standardize them.

5. Ensure all numeric values (e.g., Year Built, Size) are of the correct data type.

## Part 2: Feature Engineering and Modeling

1. Derive any useful new columns (e.g., the age of the house from the year built).

2. Prepare categorical features for use in a machine learning model by converting them into numerical form.

3. Split the dataset into training and testing parts.

4. Build and train a model to predict the price of a house.

5. Evaluate the performance of your model using appropriate metrics.

## Deliverables

1. A cleaned version of the dataset.  
2. A written explanation of the cleaning steps taken.  
3. A brief summary of the model used and its performance.  
4. Visualizations that help explain your approach or results.