# House Price Prediction using Linear Regression
# Internship Task 3 Report

## 1. Abstract

This project implements simple and multiple linear regression techniques to predict house prices based on various socio-economic and structural features. The objective is to analyze relationships among variables, evaluate model performance using statistical metrics, and derive meaningful insights for real-world real estate decision-making.

## 2. Problem Statement

To build an efficient machine learning model capable of accurately predicting housing prices using linear regression algorithms while ensuring model reliability, interpretability, and statistical robustness.

## 3. Dataset Description

The dataset contains housing-related attributes such as area, number of bedrooms, bathrooms, parking availability, furnishing status, and other amenities. The target variable is the house price.

## 4. Methodology

- Data loading and preprocessing
- Encoding of categorical variables
- Feature-target splitting
- Train-test data splitting
- Model training using Linear, Ridge, and Lasso Regression
- Model evaluation using MAE, MSE, RMSE, and $R^2$ score
- Residual diagnostics and cross-validation

## 5. Model Evaluation

The regression models were evaluated using multiple performance metrics. Cross-validation was applied to ensure stability and robustness. Residual analysis was conducted to verify randomness of errors and detect potential biases.

## 6. Feature Importance Analysis

Key features such as area, bathrooms, air conditioning, and parking were identified as the most influential factors contributing to house prices. This insight provides valuable business understanding for real estate stakeholders.

## 7. Results and Discussion

The multiple linear regression model demonstrated strong predictive capability with high $R^2$ values. Ridge and Lasso regression provided regularization benefits and improved model generalization.

## 8. Conclusion

This project successfully demonstrates a complete machine learning workflow, from preprocessing and regression modeling to advanced diagnostics and interpretation. The developed solution can be effectively used for real-world real estate price prediction and decision support.