# House Price Prediction using Linear Regression

## Project Summary

This project implements a Linear Regression model to predict the prices of houses using the House Prices: Advanced Regression Techniques dataset from Kaggle. The key features used for prediction include:  
  
- GrLivArea – Square footage of above-ground living space  
- BedroomAbvGr – Number of bedrooms above ground  
- FullBath – Number of full bathrooms above ground  
  
Using these features, a predictive model was developed and evaluated using Mean Squared Error (MSE) and R² Score. A scatter plot was also created to visually compare the actual vs predicted prices.

## Conclusion

The project successfully demonstrates the application of Linear Regression in predicting housing prices based on key numerical features. The results showed a reasonably good fit, with the predicted prices closely aligning with the actual prices for many data points.  
  
The visualization confirmed that the model captures general price trends, although some predictions deviate, indicating areas for improvement using more advanced models or feature engineering.

## Future Scope

- Add more predictive features (e.g., GarageArea, TotalBsmtSF, YearBuilt)  
- Use Polynomial Regression, Ridge, or Lasso Regression  
- Compare performance with Random Forest or Gradient Boosting  
- Build an interactive Streamlit web app to deploy the model  
- Perform cross-validation to improve generalization

## References

- Kaggle Dataset: House Prices - Advanced Regression Techniques (https://www.kaggle.com/c/house-prices-advanced-regression-techniques)  
- Scikit-learn Documentation: https://scikit-learn.org/  
- Python Libraries: pandas, numpy, matplotlib, seaborn, scikit-learn