Predicting House Prices with a Linear Regression Model

Kevin Thompson, Sterling Beason, & Brandon Croom

Data Science Program, Southern Methodist University, USA

Abstract

Price prediction is pivotal for real estate. Homeowners on the sell-side want to know when to sell, what to renovate, and how much profit they can expect from their efforts. Homebuyers want to know whether they are getting a fair price, where to look for homes in their budget, and the various trade-offs that accompany a purchasing decision. Real estate companies navigate both sides of real estate; hence, they too are a key stakeholder. In the first part of our analysis, we estimate the relationship between house prices, the square footage, and neighborhood location in Ames, Iowa. In the second part of our analysis, we train a linear regression model to predict house prices in Ames, Iowa.

1 Introduction

Ramsey and Schafer (2013)

2 Ames, Iowa Data

Kaggle (2016)

3 Analysis Question I

Ramsey and Schafer (2013)

Pearl (2009)

Ruppert and Matteson (2015)

4 Analysis Question II

Hastie et al. (2009) Trefethen and Bau (1997)

5 Appendix

References

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.

Kaggle (2016). House prices. data retrieved from the Kaggle website, https://www.kaggle.com/c/house-prices-advanced-regression-techniques/data.

Pearl, J. (2009). *Causality: Models, Reasoning and Inference*. Cambridge University Press.

Ramsey, F. and Schafer, D. (2013). *The Statistical Sleuth: A Course in Methods of Data Analysis*. Brooks/Cole Publishing Company.

Ruppert, D. and Matteson, D. (2015). *Statistics and Data Analysis for Financial Engineering*. Springer.

Trefethen, L. and Bau, D. (1997). *Numerical Linear Algebra*. Society for Industrial and Applied Mathematics.