CSC 423: Graduate Students (Online Section)
Ryan Patrick
Akbar Aidarov
Amy Aumpansub

Boom or Bust? Determinants that drive house prices

Over the last 10 - 12 years, the national housing market has fluctuated from a run up to a market bubble in 2006, to a complete collapse of the market from 2007 - 2012, and finally to a period of stabilization and recovery in recent years. Research has uncovered many of the major root causes of the boom and bust of the market. While it is important to monitor the macro indicators that might signal such risky market conditions in the future, it is also critical for both retail and commercial buyers and sellers of property to understand the micro level determinants of housing prices in order to properly assess local market conditions and risks. This analysis is a study of one such local market and the determinant variables that drive its housing prices. Through the framework of a statistical modeling approach known as the hedonic housing price model, this is a study of the housing market in Ames, Iowa from a dataset obtained from the predictive modeling competition website called Kaggle.com.

The hedonic price model is a form of a statistical tool called regression analysis which is used to estimate the degree to which one variable can be changed or impacted by other variables. By estimating these relationships, a model can then be constructed to predict the value of one variable based on changes in the other variables. The hedonic model combines structural characteristics of a property, including square footage and number of bedrooms, with location and environmental factors such as proximity to public amenities, to examine housing market prices. This study uses the hedonic price model as a framework to form hypotheses regarding which variables to include in our model for predicting prices in the Ames, Iowa housing market. With this in mind, the objectives of this analysis are threefold: to determine the most significant factors that influence housing prices, to determine if location and environmental factors cause a multiplier effect on housing prices when they are combined with structural characteristics, and to build a statistical model that reliably predicts housing prices. The outcomes are useful for both buyers and sellers in valuing a property and determining how one house compares to others within the Ames, Iowa market. Beginning with 299 possible influential variables, the process of determining the most appropriate statistical model included building and comparing multiple versions of the model to eliminate irrelevant variables that were not helpful for determining housing prices.

As a result, the final statistical model yields a reliable predictive value for housing prices along with 56 variables that explain this prediction. The top factors that are most significant in predicting housing prices are overall quality and condition of the house, the total indoor square footage, pool quality and quality of the material on the exterior. The remaining 46 variables also include location and environmental factors such as the neighborhood and proximity to local amenities. Thus, a house with greater total indoor square footage, higher pool and overall quality of the house and its conditions, etc. tends to have a higher percentage change in sale prices. Compared to location and environmental factors, the interior and exterior house features seem to have a stronger significant effect on the changes in sale prices of a house. Buyers and sellers can use this information to determine the best opportunities for improvement or investment in their properties as well as for comparison to other properties in the market.

CSC 423: Graduate Students (Online Section)
Ryan Patrick
Akbar Aidarov
Amy Aumpansub

Model Statement

ln_saleprice = 10.71827 + 0.00027213LotFrontage + 0.00000228LotArea + 0.24126numStreet1 - 0.02269numLotShape1 - 0.06321numNeighborhood1 + 0.1871numOverallQual1 + 0.20812numOverallCond1 - 0.05314numYearBuilt1 + 0.04035numYearRemodAdd1 + 0.00005507MasVnrArea - 0.2243108numExterQual1 + 0.06407numBsmtExposure1 - 0.04842numBsmtFinType21 + 0.00011982TotalHouseSF + 0.10102numCentAir1 + 0.00012855GrLivArea + 0.02698TotalBaths - 0.11625KitchenAbvGr - 0.05226numKitchenQual1 - 0.05502numFunctional1 + 0.03068Fireplaces + 0.084numGarageYrBlt1 + 0.03001GarageCars + 0.00014919GarageArea + 0.00008783WoodDeckSF + 0.00020114OpenPorchSF + 0.0031022ScreenPorch - 2.76256numPoolQC1 - 0.08211numSaleCondition1 + 0.069748numNeighborhood2 + 0.27079numOverallQual2 + 0.25519numOverallCond2 - 0.04131numYearBuilt2 - 0.08071numExterQual2 + 0.03005numFoundation2 + 0.02426numBsmtFinType12 - 0.24658numFunctional2 - 0.0218numGarageFinish2 + 0.07294numExterior1st3 + 0.03598numMasVnrType3 - 0.1118numExterQual3 - 0.33715numHeating3 - 0.04221numGarageYrBlt3 - 0.04968numGarageQual3 + 0.06853numSaleType3 - 0.10021numMSZoning4 + 0.06725numYearRemodAdd4 - 0.03354numHeatingQC4 + 0.27903numGarageQual4 + 0.05513numYearBuilt5 + 0.09093numYearRemodAdd5 - 0.10946numRoofMatl5 - 0.21626numFoundation5 - 0.03221numBsmtFinType16 + 0.178numExterior1st10 + 0.08776numExterior1st11 + e.