Exploring the effect of different aspects of residential homes on price and predicting the price of houses in Ames, Iowa

1. Problem description

Most people understand and accept the risk that stock prices might fall, but when they buy a house, they might never think that the price of house will decrease. The housing market is subject to the same forces of supply and demand as any other market. There are more factors that affect the house price rather than the number of bedrooms or a white picket fence. In this paper, we explore almost every aspects of house in determining the house price. We perform the linear regression techniques to predict the house price, evaluate models and choose the best regression model.

2. Data analysis

The Ames housing dataset is taken from Kaggle. The specific URL is <https://www.kaggle.com/c/house-prices-advanced-regression-techniques>. There are 79 house features and 1460 observations. These 79 variables are cover almost every aspects of house. So this dataset is a good to use for house price predicting.

SalePrice - the property's sale price in dollars. This is the target variable that you're trying to predict.

1. MSSubClass: The building class
2. MSZoning: The general zoning classification
3. LotFrontage: Linear feet of street connected to property
4. LotArea: Lot size in square feet
5. Street: Type of road access
6. Alley: Type of alley access
7. LotShape: General shape of property
8. LandContour: Flatness of the property
9. Utilities: Type of utilities available
10. LotConfig: Lot configuration
11. LandSlope: Slope of property
12. Neighborhood: Physical locations within Ames city limits
13. Condition1: Proximity to main road or railroad
14. Condition2: Proximity to main road or railroad (if a second is present)
15. BldgType: Type of dwelling
16. HouseStyle: Style of dwelling
17. OverallQual: Overall material and finish quality
18. OverallCond: Overall condition rating
19. YearBuilt: Original construction date
20. YearRemodAdd: Remodel date
21. RoofStyle: Type of roof
22. RoofMatl: Roof material
23. Exterior1st: Exterior covering on house
24. Exterior2nd: Exterior covering on house (if more than one material)
25. MasVnrType: Masonry veneer type
26. MasVnrArea: Masonry veneer area in square feet
27. ExterQual: Exterior material quality
28. ExterCond: Present condition of the material on the exterior
29. Foundation: Type of foundation
30. BsmtQual: Height of the basement
31. BsmtCond: General condition of the basement
32. BsmtExposure: Walkout or garden level basement walls
33. BsmtFinType1: Quality of basement finished area
34. BsmtFinSF1: Type 1 finished square feet
35. BsmtFinType2: Quality of second finished area (if present)
36. BsmtFinSF2: Type 2 finished square feet
37. BsmtUnfSF: Unfinished square feet of basement area
38. TotalBsmtSF: Total square feet of basement area
39. Heating: Type of heating
40. HeatingQC: Heating quality and condition
41. CentralAir: Central air conditioning
42. Electrical: Electrical system
43. 1stFlrSF: First Floor square feet
44. 2ndFlrSF: Second floor square feet
45. LowQualFinSF: Low quality finished square feet (all floors)
46. GrLivArea: Above grade (ground) living area square feet
47. BsmtFullBath: Basement full bathrooms
48. BsmtHalfBath: Basement half bathrooms
49. FullBath: Full bathrooms above grade
50. HalfBath: Half baths above grade
51. Bedroom: Number of bedrooms above basement level
52. Kitchen: Number of kitchens
53. KitchenQual: Kitchen quality
54. TotRmsAbvGrd: Total rooms above grade (does not include bathrooms)
55. Functional: Home functionality rating
56. Fireplaces: Number of fireplaces
57. FireplaceQu: Fireplace quality
58. GarageType: Garage location
59. GarageYrBlt: Year garage was built
60. GarageFinish: Interior finish of the garage
61. GarageCars: Size of garage in car capacity
62. GarageArea: Size of garage in square feet
63. GarageQual: Garage quality
64. GarageCond: Garage condition
65. PavedDrive: Paved driveway
66. WoodDeckSF: Wood deck area in square feet
67. OpenPorchSF: Open porch area in square feet
68. EnclosedPorch: Enclosed porch area in square feet
69. 3SsnPorch: Three season porch area in square feet
70. ScreenPorch: Screen porch area in square feet
71. PoolArea: Pool area in square feet
72. PoolQC: Pool quality
73. Fence: Fence quality
74. MiscFeature: Miscellaneous feature not covered in other categories
75. MiscVal: $Value of miscellaneous feature
76. MoSold: Month Sold
77. YrSold: Year Sold
78. SaleType: Type of sale
79. SaleCondition: Condition of sale