

**HOUSING: PRICE PREDICTION**

Submitted by:

SHRUTI KUMARI

**ACKNOWLEDGMENT**

-Machine learning recorded videos of Data Trained

-Study few research paper of Real estate Industry

-data description details mostly search on Google and understand

-some websites used: kaggle, geeksforgeeks, w3resources, towardsdatascience

**INTRODUCTION**

Real Estate is the least transparent industry in our ecosystem. Housing prices keep changing day by day and sometimes are hyped rather than being based on valuation. Predicting house prices with real factors is the main crux of my project. Here I aim to make our evaluation based on every basic parameter that is considered while determining the price. I am using best regression techniques in this house price prediction.

* Hardware and Software Requirements and Tools Used

**-**Python

-numpy and pandas library

-matplotlib and seaborn for plotting

-Label encoding for encoding the data

-RandomForestRegressor library for regression of the data

**Model/s Development and Evaluation**

-House Price Prediction Project has two data as train.csv and test.csv

-Importing the important libraries

-load the data through pandas as pd and check the shape

-through pandas set options to display max column and rows

-Join train and test data through pd.concat

-1460 rows and 81 columns in this data

-There are many null values in this dataset  
-There is a library ‘sweetviz’ that visualize each data with their description like max, missing values, min etc.

-I divide the whole data in categorical and numerical list

* Categorical: ['MSZoning', 'Street', 'Alley', 'LotShape', 'LandContour', 'Utilities', 'LotConfig', 'LandSlope', 'Neighborhood', 'Condition1', 'Condition2', 'BldgType', 'HouseStyle', 'RoofStyle', 'RoofMatl', 'Exterior1st', 'Exterior2nd', 'MasVnrType', 'ExterQual', 'ExterCond', 'Foundation','BsmtQual', 'BsmtCond', 'BsmtExposure', 'BsmtFinType1', 'BsmtFinType2','Heating', 'HeatingQC', 'CentralAir', 'Electrical', 'KitchenQual','Functional', 'FireplaceQu', 'GarageType', 'GarageFinish', 'GarageQual','GarageCond', 'PavedDrive', 'PoolQC', 'Fence', 'MiscFeature','SaleType', 'SaleCondition']
* Numerical: ['Id', 'MSSubClass', 'LotFrontage', 'LotArea', 'OverallQual',

'OverallCond', 'YearBuilt', 'YearRemodAdd', 'MasVnrArea', 'BsmtFinSF1',

'BsmtFinSF2', 'BsmtUnfSF', 'TotalBsmtSF', '1stFlrSF', '2ndFlrSF',

'LowQualFinSF', 'GrLivArea', 'BsmtFullBath', 'BsmtHalfBath', 'FullBath',

'HalfBath', 'BedroomAbvGr', 'KitchenAbvGr', 'TotRmsAbvGrd',

'Fireplaces', 'GarageYrBlt', 'GarageCars', 'GarageArea', 'WoodDeckSF',

'OpenPorchSF', 'EnclosedPorch', '3SsnPorch', 'ScreenPorch', 'PoolArea',

'MiscVal', 'MoSold', 'YrSold', 'SalePrice']

-Visualize all the categorical data using histogram and numerical data using scatter plot

-plot the target variable sale price

-delete the columns that have more than 80% of null values

-fill the numerical null values with mean and categorical null values with mode.

-use one hot encoder

- use traintestsplit to train the 80% data and test 20% data

-use RandomForestRegressor to predict the sale price of testing data

**CONCLUSION**

-Test data id with sale price

| **Id** | **SalePrice** |
| --- | --- |
| **0** | 337 | 126277.547163 |
| **1** | 1018 | 198213.819894 |
| **2** | 929 | 138814.620755 |
| **3** | 1148 | 132851.337342 |
| **4** | 1227 | 138256.173259 |
| **5** | 650 | 160673.276911 |
| **6** | 1453 | 156950.321882 |
| **7** | 152 | 226083.824755 |
| **8** | 427 | 154333.781390 |
| **9** | 776 | 420751.429067 |
| **10** | 30 | 147371.146767 |
| **11** | 1425 | 195473.843570 |
| **12** | 423 | 125651.789127 |
| **13** | 1185 | 280811.130372 |
| **14** | 775 | 112955.439510 |
| **15** | 391 | 146116.140086 |
| **16** | 1408 | 144354.576563 |
| **17** | 513 | 225753.173318 |
| **18** | 1266 | 194697.633019 |
| **19** | 173 | 124807.978647 |
| **20** | 1150 | 155310.072529 |
| **21** | 797 | 121352.768995 |
| **22** | 137 | 145522.503438 |
| **23** | 706 | 148736.282805 |
| **24** | 1377 | 142710.520649 |
| **25** | 1177 | 223966.447318 |
| **26** | 1456 | 212904.511187 |
| **27** | 369 | 156817.831690 |
| **28** | 1421 | 117836.522396 |
| **29** | 999 | 313933.318443 |
| **30** | 1217 | 153522.309127 |
| **31** | 937 | 140799.909223 |
| **32** | 769 | 277626.409450 |
| **33** | 831 | 263229.615906 |
| **34** | 678 | 144489.634647 |
| **35** | 574 | 197231.548815 |
| **36** | 921 | 210243.934611 |
| **37** | 1292 | 117880.913007 |
| **38** | 1277 | 149482.161342 |
| **39** | 676 | 172957.070193 |
| **40** | 108 | 141185.842781 |
| **41** | 424 | 307271.932575 |
| **42** | 823 | 161787.440695 |
| **43** | 1455 | 355871.242623 |
| **44** | 377 | 158860.824685 |
| **45** | 1256 | 195245.285402 |
| **46** | 1120 | 149551.506683 |
| **47** | 265 | 158824.899918 |
| **48** | 1158 | 149888.187450 |
| **49** | 725 | 146738.867822 |
| **50** | 1378 | 249264.336563 |
| **51** | 1139 | 255447.794755 |
| **52** | 122 | 163009.321271 |
| **53** | 515 | 172333.262911 |
| **54** | 518 | 194210.873318 |
| **55** | 1214 | 198425.466889 |
| **56** | 443 | 147406.665043 |
| **57** | 903 | 161180.569390 |
| **58** | 53 | 146199.851378 |
| **59** | 469 | 142080.867474 |
| **60** | 76 | 170667.537582 |
| **61** | 1142 | 143835.669390 |
| **62** | 1222 | 120186.864947 |
| **63** | 851 | 161124.516697 |
| **64** | 334 | 186360.021630 |
| **65** | 1008 | 279793.671127 |
| **66** | 19 | 259556.741235 |
| **67** | 339 | 215787.193762 |
| **68** | 1118 | 122253.926875 |
| **69** | 834 | 156564.753762 |
| **70** | 1176 | 109005.420695 |
| **71** | 945 | 180260.045570 |
| **72** | 1415 | 178633.962733 |
| **73** | 911 | 125854.820743 |
| **74** | 388 | 164896.139786 |
| **75** | 747 | 248237.230575 |
| **76** | 620 | 215438.280611 |
| **77** | 386 | 111220.695283 |
| **78** | 818 | 257272.861079 |
| **79** | 539 | 129195.024935 |
| **80** | 535 | 169607.999295 |
| **81** | 1399 | 232238.543247 |
| **82** | 571 | 360796.212779 |
| **83** | 38 | 154480.129702 |
| **84** | 850 | 147092.379894 |
| **85** | 1032 | 175760.831259 |
| **86** | 1433 | 132941.231330 |
| **87** | 1354 | 235212.461163 |
| **88** | 1072 | 142223.861618 |
| **89** | 371 | 161921.047726 |
| **90** | 1242 | 161377.419271 |
| **91** | 681 | 139660.181414 |
| **92** | 290 | 145960.746086 |
| **93** | 973 | 280360.062563 |
| **94** | 989 | 210860.771858 |
| **95** | 484 | 253215.220563 |
| **96** | 1240 | 257841.884887 |
| **97** | 1125 | 332880.361354 |
| **98** | 1143 | 129272.396887 |
| **99** | 1340 | 141028.481019 |
| **100** | 1343 | 206377.935366 |
| **101** | 936 | 254223.557031 |
| **102** | 1151 | 115024.581594 |
| **103** | 1380 | 202573.989199 |
| **104** | 1190 | 179118.912949 |
| **105** | 635 | 265961.480575 |
| **106** | 47 | 178507.651151 |
| **107** | 729 | 149620.829462 |
| **108** | 1434 | 200771.918455 |
| **109** | 472 | 156234.168911 |
| **110** | 1156 | 183566.780707 |
| **111** | 1352 | 166238.278146 |
| **112** | 717 | 139833.866923 |
| **113** | 385 | 102609.930443 |
| **114** | 1334 | 151159.097307 |
| **115** | 243 | 155402.595031 |
| **116** | 39 | 184440.863798 |
| **117** | 1168 | 171897.042983 |
| **118** | 214 | 124415.476611 |
| **119** | 647 | 118454.598791 |
| **120** | 490 | 157861.961450 |
| **121** | 512 | 147437.685247 |
| **122** | 1181 | 126513.370995 |
| **123** | 1451 | 138451.511330 |
| **124** | 1428 | 119355.496779 |
| **125** | 767 | 246295.058743 |
| **126** | 1250 | 155262.505834 |
| **127** | 855 | 150910.830743 |
| **128** | 1001 | 290448.528515 |
| **129** | 49 | 148588.574971 |
| **130** | 25 | 188508.616325 |
| **131** | 1058 | 148341.090563 |
| **132** | 24 | 251092.348671 |
| **133** | 758 | 131769.986815 |
| **134** | 1060 | 134078.737163 |
| **135** | 1110 | 186210.226995 |
| **136** | 1057 | 196823.003223 |
| **137** | 491 | 218313.273211 |
| **138** | 378 | 165130.992086 |
| **139** | 1429 | 154914.841175 |
| **140** | 55 | 141608.297654 |
| **141** | 770 | 196701.389318 |
| **142** | 737 | 138781.581103 |
| **143** | 59 | 147669.381570 |
| **144** | 29 | 170757.653438 |
| **145** | 788 | 345055.685342 |
| **146** | 9 | 310528.048192 |
| **147** | 1371 | 160671.039163 |
| **148** | 260 | 208484.855091 |
| **149** | 249 | 145963.067498 |
| **150** | 88 | 337571.477786 |
| **151** | 893 | 112843.217318 |
| **152** | 803 | 208685.499127 |
| **153** | 901 | 124825.400887 |
| **154** | 1124 | 136409.393019 |
| **155** | 223 | 178215.920122 |
| **156** | 1278 | 121184.008983 |
| **157** | 1174 | 155641.137474 |
| **158** | 312 | 170406.307774 |
| **159** | 624 | 192366.754396 |
| **160** | 1445 | 156143.861438 |
| **161** | 296 | 155735.853606 |
| **162** | 753 | 161782.459402 |
| **163** | 1320 | 205269.422839 |
| **164** | 432 | 396654.750336 |
| **165** | 1362 | 194535.213414 |
| **166** | 1069 | 224154.723247 |
| **167** | 1255 | 110158.012539 |
| **168** | 50 | 240577.005330 |
| **169** | 454 | 252409.959295 |
| **170** | 881 | 182678.345223 |
| **171** | 500 | 183473.580887 |
| **172** | 807 | 138743.730527 |
| **173** | 766 | 227280.642887 |
| **174** | 866 | 178669.680719 |
| **175** | 162 | 119960.025007 |
| **176** | 1193 | 160035.307654 |
| **177** | 798 | 137975.205295 |
| **178** | 1400 | 145473.353115 |
| **179** | 6 | 236353.683115 |
| **180** | 58 | 179810.698385 |
| **181** | 1435 | 203542.016791 |
| **182** | 867 | 156664.122875 |
| **183** | 960 | 141438.789750 |
| **184** | 441 | 185798.539140 |
| **185** | 481 | 184019.591187 |
| **186** | 219 | 177226.302481 |
| **187** | 128 | 137930.795127 |
| **188** | 858 | 163195.769798 |
| **189** | 215 | 141001.776851 |
| **190** | 1030 | 165746.419786 |
| **191** | 1204 | 145247.968050 |
| **192** | 531 | 123286.340731 |
| **193** | 529 | 199076.161942 |
| **194** | 346 | 157478.834026 |
| **195** | 621 | 166988.680983 |
| **196** | 1194 | 176670.651354 |
| **197** | 1000 | 132972.826887 |
| **198** | 761 | 252133.874599 |
| **199** | 1351 | 135828.137139 |
| **200** | 1442 | 333160.040515 |
| **201** | 233 | 132147.243043 |
| **202** | 360 | 213725.100455 |
| **203** | 886 | 142867.155079 |
| **204** | 11 | 136303.523259 |
| **205** | 572 | 153714.939690 |
| **206** | 1375 | 186042.691486 |
| **207** | 328 | 126732.748182 |
| **208** | 1417 | 175720.993522 |
| **209** | 1089 | 131074.725043 |
| **210** | 242 | 292066.302336 |
| **211** | 1198 | 174826.741426 |
| **212** | 492 | 257344.542671 |
| **213** | 743 | 186367.575199 |
| **214** | 1258 | 288521.369378 |
| **215** | 1350 | 164010.830122 |
| **216** | 722 | 142365.938433 |
| **217** | 442 | 248718.548803 |
| **218** | 90 | 139033.558995 |
| **219** | 1007 | 184720.041247 |
| **220** | 1241 | 141818.681342 |
| **221** | 144 | 141175.085295 |
| **222** | 1317 | 144207.355486 |
| **223** | 1080 | 148536.759762 |
| **224** | 1160 | 180115.201474 |
| **225** | 482 | 164824.815918 |
| **226** | 33 | 127449.432839 |
| **227** | 1280 | 169515.396721 |
| **228** | 541 | 173889.926469 |
| **229** | 994 | 182027.328146 |
| **230** | 1385 | 142564.761690 |
| **231** | 792 | 187183.891942 |
| **232** | 1002 | 351769.628515 |
| **233** | 420 | 198223.511139 |
| **234** | 354 | 108623.096899 |
| **235** | 447 | 201125.031271 |
| **236** | 883 | 223592.378539 |
| **237** | 578 | 119890.100911 |
| **238** | 1251 | 127082.006803 |
| **239** | 671 | 173901.224481 |
| **240** | 1017 | 172523.293618 |
| **241** | 780 | 180090.472423 |
| **242** | 1269 | 89203.436671 |
| **243** | 1145 | 160284.557546 |
| **244** | 486 | 177511.054901 |
| **245** | 134 | 154052.260217 |
| **246** | 1282 | 96692.806587 |
| **247** | 693 | 136486.412995 |
| **248** | 1119 | 231897.915151 |
| **249** | 408 | 123561.527510 |
| **250** | 1288 | 170428.749031 |
| **251** | 159 | 248885.062599 |
| **252** | 577 | 126706.090983 |
| **253** | 1014 | 164412.607546 |
| **254** | 226 | 167725.777187 |
| **255** | 140 | 220918.378959 |
| **256** | 109 | 212638.904731 |
| **257** | 988 | 221842.560433 |
| **258** | 633 | 194472.518205 |
| **259** | 1221 | 113222.755295 |
| **260** | 1431 | 151906.339055 |
| **261** | 365 | 203398.869486 |
| **262** | 495 | 264049.805462 |
| **263** | 165 | 142977.237009 |
| **264** | 816 | 283804.408719 |
| **265** | 181 | 143062.449474 |
| **266** | 107 | 120323.125390 |
| **267** | 125 | 205138.804433 |
| **268** | 1218 | 127199.064815 |
| **269** | 938 | 160054.623223 |
| **270** | 372 | 145579.201139 |
| **271** | 45 | 132943.921474 |
| **272** | 995 | 185425.706709 |
| **273** | 314 | 144619.377366 |
| **274** | 948 | 172967.428923 |
| **275** | 217 | 144743.261762 |
| **276** | 781 | 322460.449199 |
| **277** | 1293 | 129971.200587 |
| **278** | 1005 | 186332.950313 |
| **279** | 528 | 129343.316647 |
| **280** | 44 | 186786.699740 |
| **281** | 645 | 113823.023642 |
| **282** | 340 | 219368.781019 |
| **283** | 1116 | 204353.431067 |
| **284** | 358 | 184152.369510 |
| **285** | 72 | 222233.110659 |
| **286** | 56 | 158054.943031 |
| **287** | 83 | 162293.675079 |
| **288** | 1048 | 145826.431342 |
| **289** | 17 | 247323.778767 |
| **290** | 523 | 295152.734803 |
| **291** | 1379 | 213355.809115 |