
**Recommendations on how
can home owners increase
the sale prices of their
homes?**

Russell Foo

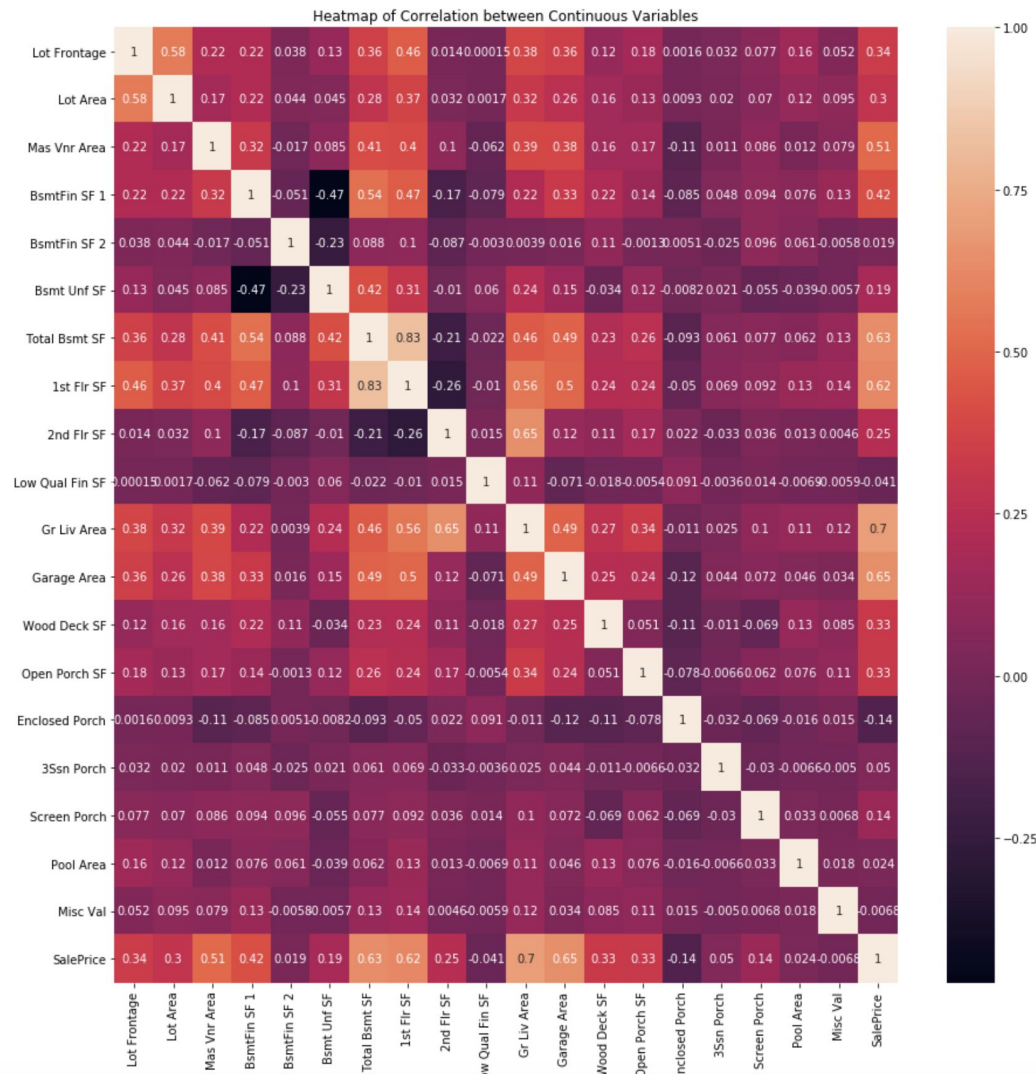
Exploratory Data Analysis

Checking for collinearity between continuous variables.

'Total Bsmt SF' vs '1st Flr SF' has a correlation of 0.83

'2nd Flr SF' vs 'Gr Liv Area' has a correlation of 0.65

Drop the features in the pair with lower correlation vs 'SalePrice'. 'Total Bsmt SF' (0.63) vs '1st Flr SF' (0.62) and '2nd Flr SF' (0.25) vs 'Gr Liv Area'(0.7). We will drop '1st Flr SF' and 'Gr Liv Area'.

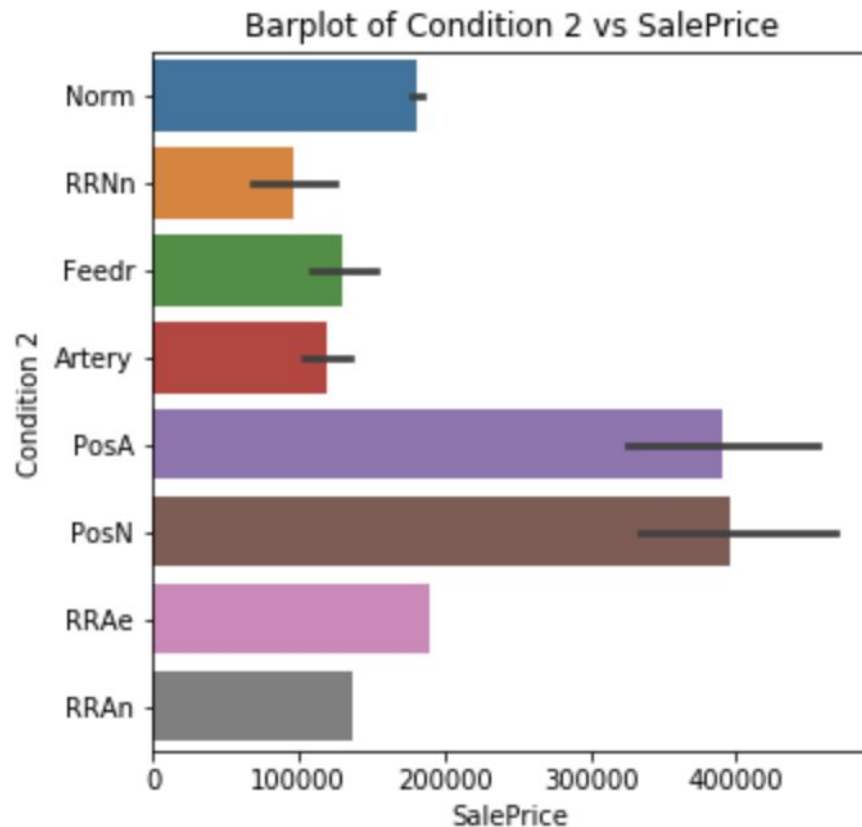


Exploratory Data Analysis

Plotting barplots between all categorical data vs SalePrice tells us which features are skewed

Values are mostly around PosA and PosN

Drop Condition 2



Exploratory Data Analysis

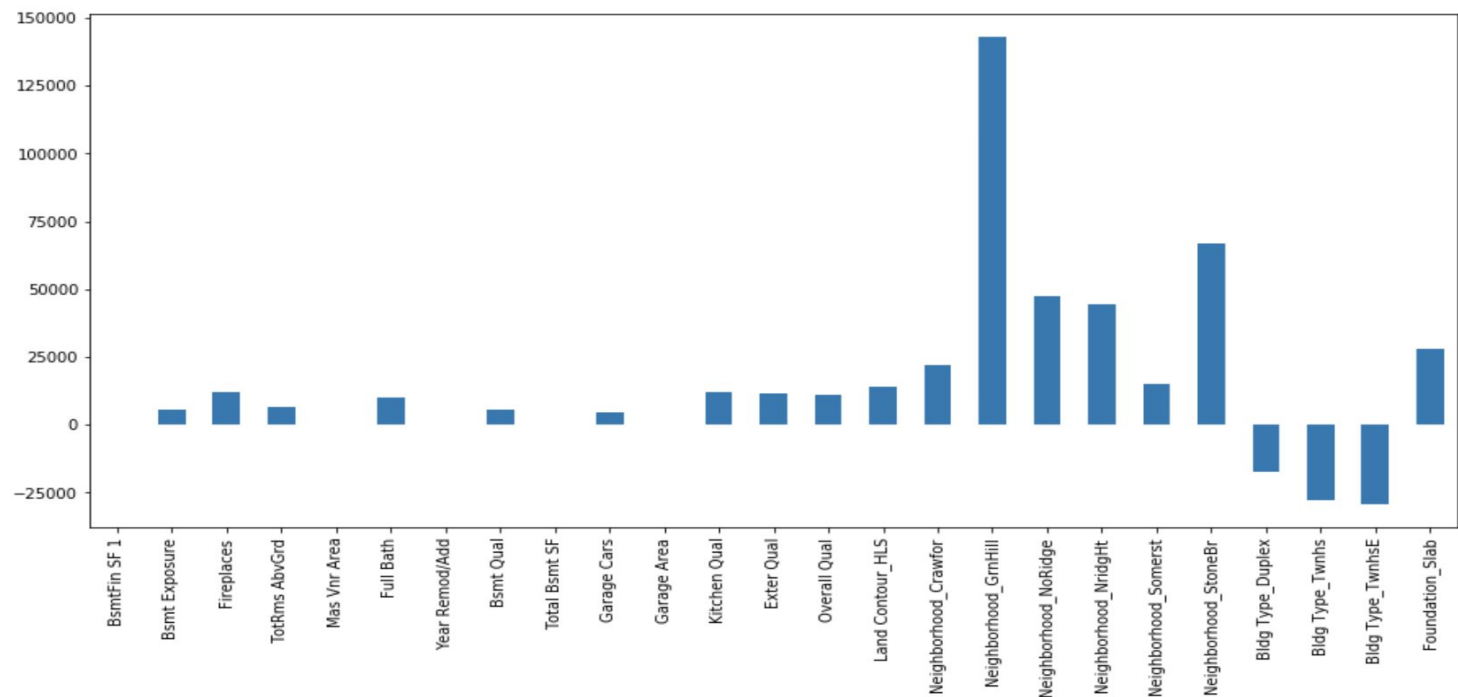
We will select numerical data from those having high correlation with our SalePrice.

```
# selecting correlation coefficients against 'SalePrice' which are about less than -0.4 and more than 0.4  
train_ord.corr()[ (train_ord.corr()[ 'SalePrice' ] <= -0.4) | (train_ord.corr()[ 'SalePrice' ] >= 0.4) ][ 'SalePrice' ].sort_values(ascending = True)
```

BsmtFin SF 1	0.424380
Bsmt Exposure	0.426241
Heating QC	0.458396
Fireplaces	0.471222
TotRms AbvGrd	0.505188
Mas Vnr Area	0.512699
Fireplace Qu	0.537287
Full Bath	0.538471
Year Remod/Add	0.549397
Garage Finish	0.558269
Year Built	0.571860
Bsmt Qual	0.615154
Total Bsmt SF	0.631975
Garage Cars	0.647015
Garage Area	0.648661
Kitchen Qual	0.691030
Exter Qual	0.711814
Overall Qual	0.799978
SalePrice	1.000000

Name: SalePrice, dtype: float64

Insights and Recommendations



Insights and Recommendations

Prices most is dependant on whether the house is located in the Neighborhood of Green Hills.

Prices hurts most if the house is if the house is a Townhouse end unit.

Features home owners should look at to improve their house prices are the number of full bathrooms above grade, Kitchen quality, Exterior quality, Overall quality, having Hill side land contour and having slab foundation.

Neighborhoods which may be considered good investments are (in order) Green Hills, Stone Brook, Northridge, Northridge Heights, Crawford and Somerset.

Improvements can be made to generalise the model if we had more data:

- Crime rates
- Distance from CBD
- Quality of Air