

Determining Housing Prices with Data Science

Real Estate Appraisers Coalition
Of Ames Iowa. *

Presentation by : CM APRIL

*(disclaimer - fictionals for funsies.)



Problem Statement

- How much is a house worth?
- What factors really matter in determining a fair market value?



How do we find what matters most?

We want to create a simplified model to assess the value of numerous homes.

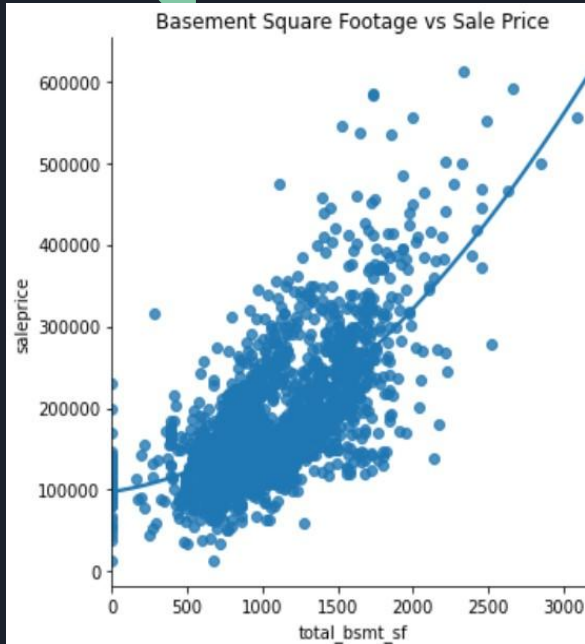
We can look at Coefficients to determine what is important.

Coefficients, and why they matter!

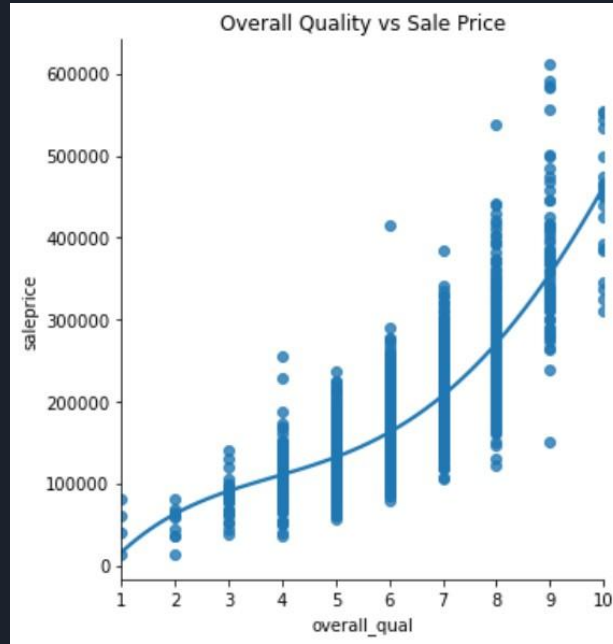
	saleprice		
saleprice	1.000000		
overall_qual	0.804410	garage_finish	0.559397
gr_liv_area	0.720759	year_remod/add	0.550784
exter_qual	0.715866	fireplace_qu	0.539121
kitchen_qual	0.694008	full_bath	0.537682
total_bsmt_sf	0.667386	garage_yr_blt	0.518043
garage_area	0.655361	mas_vnr_area	0.517000
1st_flr_sf	0.651941	totrms_abvgprd	0.509194
neighborhood	0.649567	fireplaces	0.473137
garage_cars	0.648739	heating_qc	0.459751
bsmt_qual	0.618223	bsmtfin_sf_1	0.446503
year_built	0.572934	bsmt_exposure	0.427509

- Correlation coefficients indicate the **strength of the relationship** between two different variables.
- Higher absolute value = Stronger relationship
- Most significant variables matter more to our model.

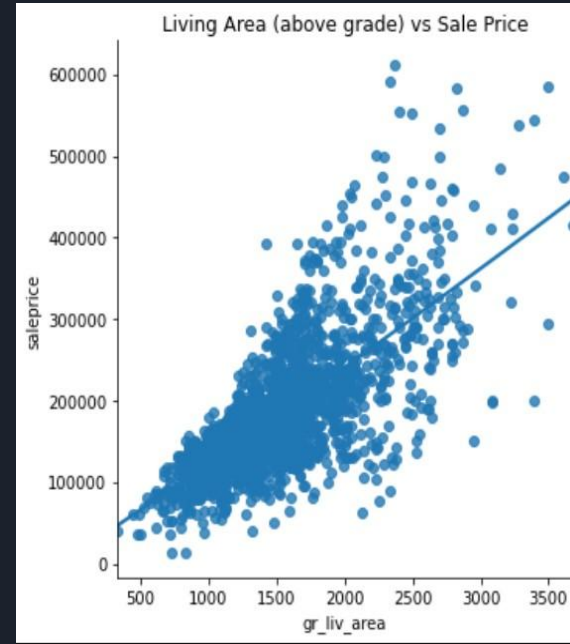
Plots of High Importance Variables



total_bsmt_sf 0.667386

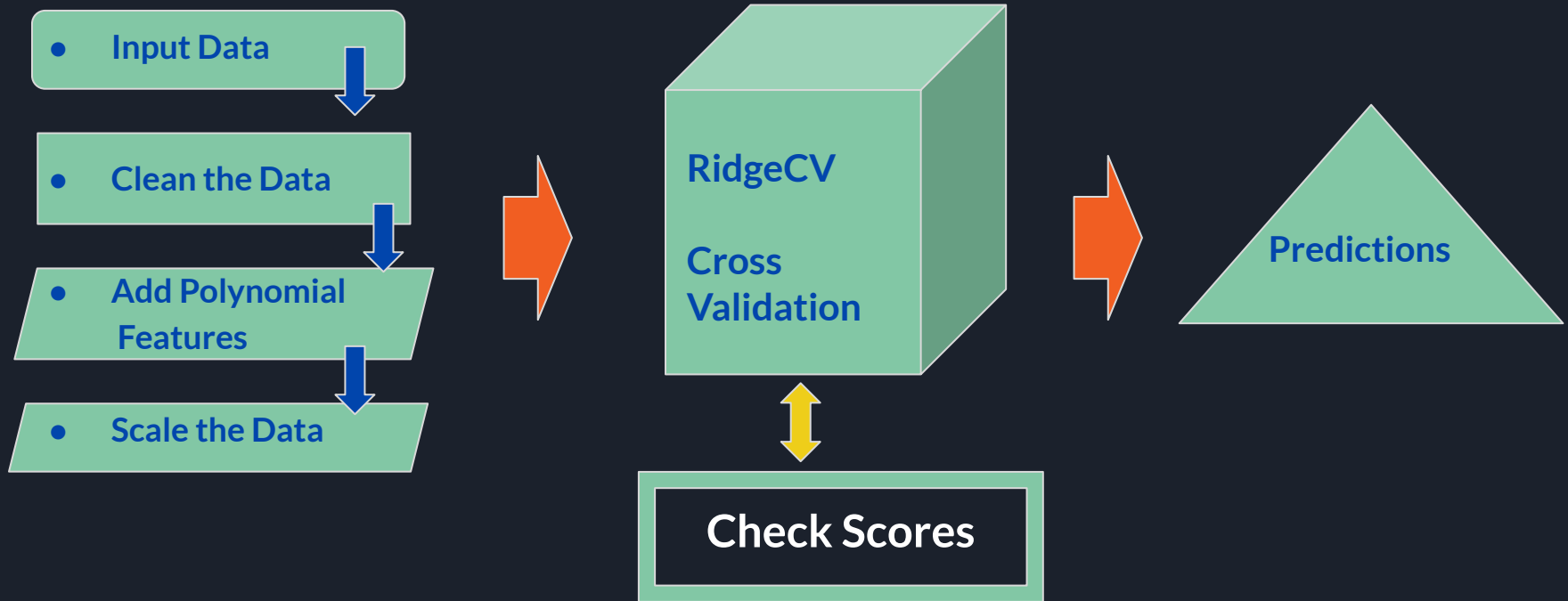


overall_qual 0.804410



gr_liv_area 0.720759

Model I fitted and how it works





Conclusions

- The model at its current development gives us predictions of 87.5% Accuracy.
- Includes 16 of the most important variables

```
print(model.score(Z_train, y_train))  
print(model.score(Z_test, y_test))
```

```
0.9383960455749727  
0.8746692493633007
```