COMO GROUP

Real Estate solutions Using Machine Learning

CHIBUZO UGONABO

WHAT MAKES A HOUSE VALUABLE?

8 Critical factors that influence home value

- Neighborhood comparisons
- Location
- Home size and usable space
- Age and condition
- Upgrades and updates
- The local market
- Economic indicators
- Interest rates

COMO GROUP

THE GOAL

The goal: Build a model to predict sale prices of houses using their features as means to Determine Investment opportunities

Result: Using Machine Learning Algorithms (linear regression), I was able build a model to predict Sale Prices of homes with a **+/- 11.69%** error margin

Data:

Kaggle: The data set data set describes homes with 79 explanatory variables describing (almost) every aspect of residential homes in Ames, Iowa, sold between 2006- 2010

COMO GROUP

WHY ?

- The real estate market is highly volatile.
- It's a buyer's market exposed to high fluctuations with no single source of truth to determine the value of a property at a given time.
- Online listings do not always display the updated/ accurate housing prices - prices are sometimes set below market value
- This model provides an actual price value to insight investment opportunities.

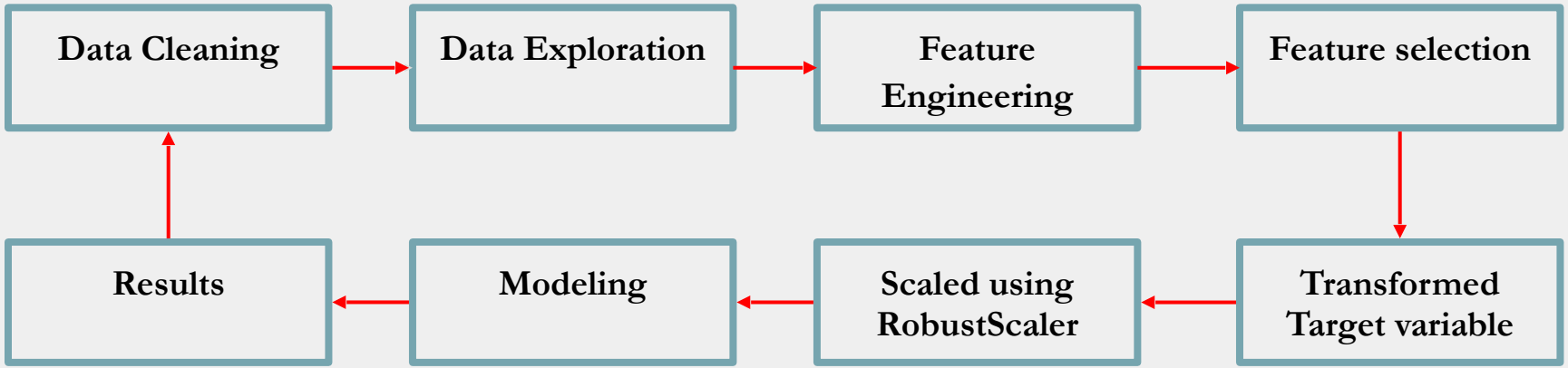
COMO GROUP

BUSINESS APPLICATION

For Investors: A source of truth to determine a good investment opportunity based on the past sale prices as it relates to features of the house. These features include the physical characteristics of the house and economic standings of its location.

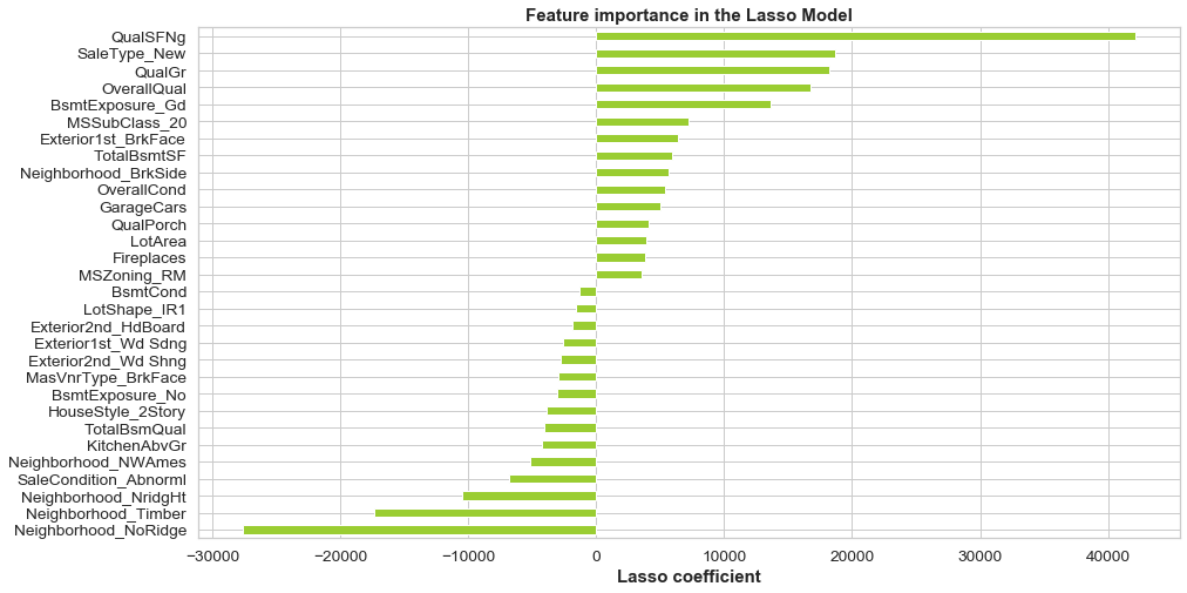
COMO GROUP

The Process



COMO GROUP

FINDINGS



I determined that the most important Feature in predicting the sale price was:

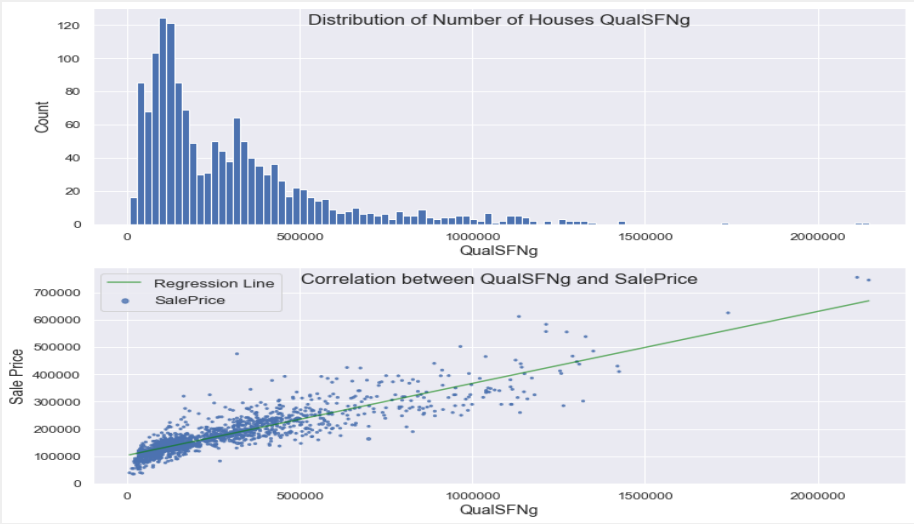
Quality- Square-Footage-Neighborhood "QualSFNg"(engineered):

A combination of Above Ground Living area Square feet + TOTAL Quality + Total Condition *Neighborhood Rating(engineered)

COMO GROUP

SO WHERE IS THE VALUE?

- **QUALITY** - Material used to construct (rated in data 1-10)
- **CONDITION**- current wear and tear (rated in data 1-10)
- **NEIGHBORHOOD RATING** - Determined by Average household income
- **LIVING AREA**- square footage



COMO GROUP

RESULTS

	Model Name	Train RMSE	Test RMSE	Test Std	Time
4	XGBRegressor	0.014955	0.116906	0.017282	84.600097
3	GradientBoostingRegressor	0.074671	0.117737	0.021712	45.165828
2	Elasticnet	0.105374	0.120443	0.017034	7.972505
1	Lasso	0.105253	0.120449	0.017268	2.815425
7	TweedieRegressor	0.104996	0.121317	0.017773	1.457334
5	SVR	0.109837	0.121354	0.023765	2.518428
0	Ridge	0.107951	0.122394	0.019966	2.875212
6	HistGradientBoostingRegressor	0.094762	0.125324	0.017231	3.335385

Conclusion:

With the **XGBR model**, an investor can predict property values at a +/- 11.69% error margin. As long as the property is bought below this margin, an investor stands to make a profit and can determine whether or not it'll be a good investment.

COMO GROUP

NEXT STEPS

- Handle the **outliers** accordingly: given the business idea, I would like to limit the model's errors by redefining the range of predictions (testing in my data was not conclusive)
- **More data** + Current data
- Creating a **pipeline** that takes in data (listings...) and recommends profitable investment opportunities
- Adding a **time** series component
- Using **alternative data** as an addition tool to get an edge on investment

COMO GROUP

THANK YOU!

COMO GROUP