Your House is Worth More than You Think!

Ground Assessment (GA)

Background

We are the **Data Science team of Ground Assessment (GA)**, a company that provides **one-stop consultation and construction services** to property owners in the United States.

Problem Statement

Recommend types of renovations/enhancements ('features') that would help property owners increase the sale price of their house.

Audience

GA management

Methodology

- Dataset: Sale price of different houses in Ames, Iowa between 2006-2010
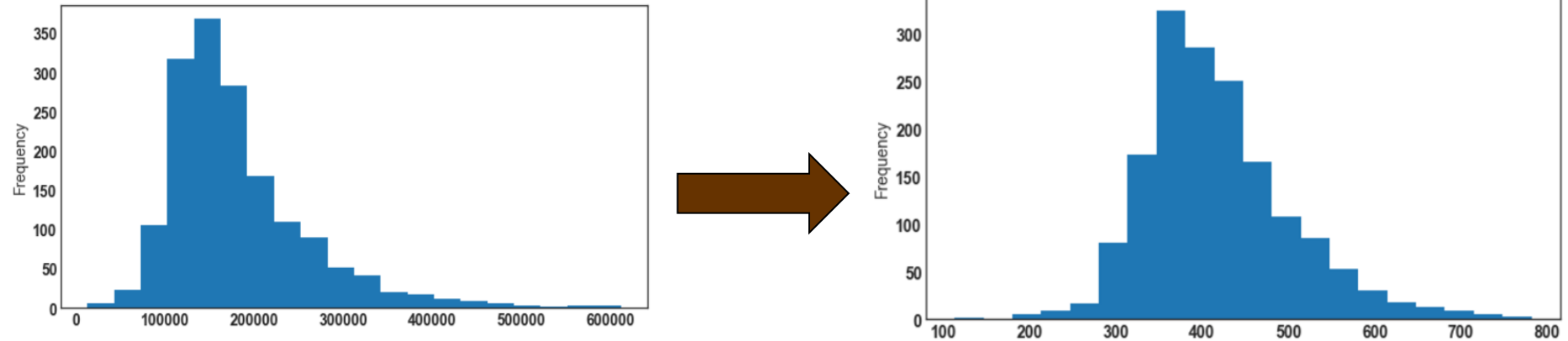
Data cleaning & exploration	Prescriptive model fitting	Top feature identification & return value estimation

- Noteworthy edits:
  - Removed 2 outliers: large houses (>5,000 sqft) sold at very low prices
  - Corrected 1 record where a garage is built in the future (in 2207)
  - Imputed missing values based on analysis and inference
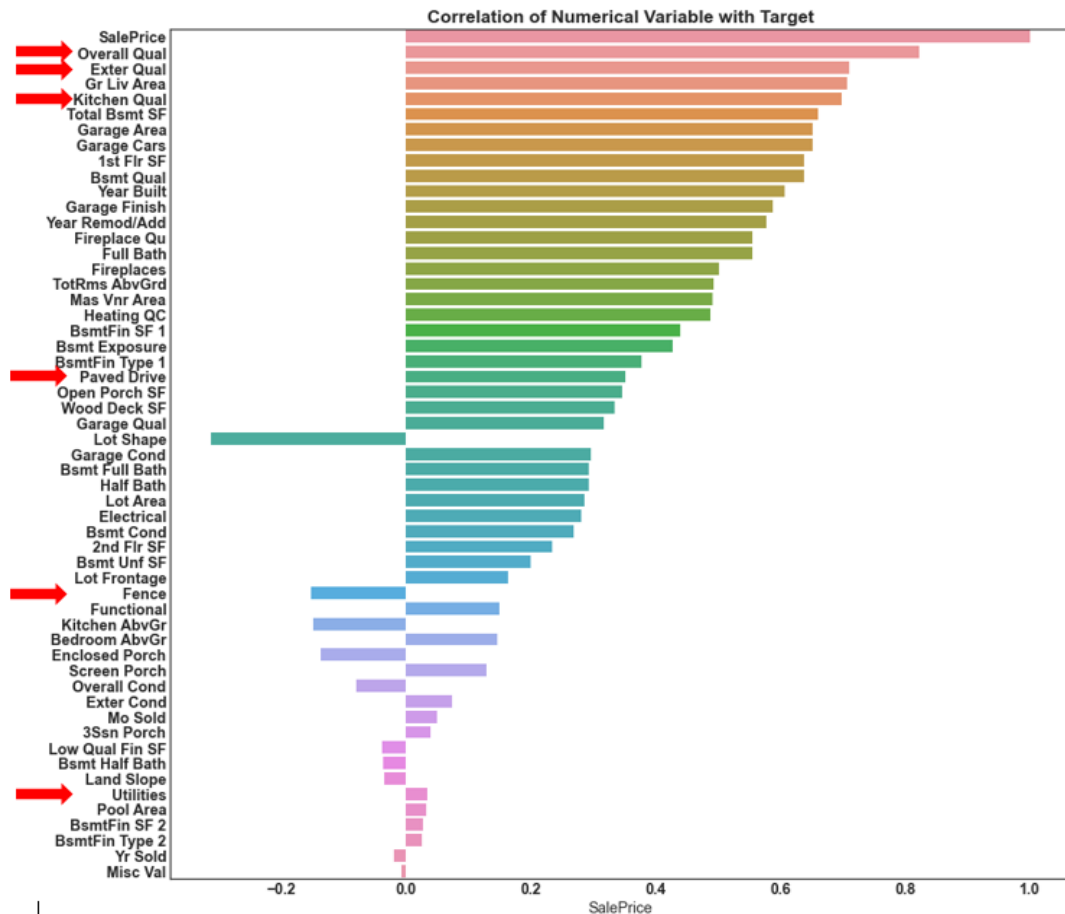
Distribution of Housing Sale Price

Positive Skewed transformed to Normal Distribution for better model

Our Prediction against Actual Sale Price

Predicted vs Actual Sale Price based on Linear Regression Model with Lasso Regularization

Feature Correlations to Housing Price

Correlation of Numerical Variable with Target

Features House Owners Can Improve On

| Feature	Description	Correlation to Price
Overall Quality	Finishing material quality	0.8
Exterior Quality	Finishing exterior quality	0.7
Kitchen Quality	Kitchen setup	0.68
Pave Drive	Top quality: Paved Lowest quality: Dirt/Gravel	0.38
Fence	Top quality: Good fencing material with privacy Lowest quality: No fencing	-0.18
Utility	Top quality: All public utilities (Electricity, gas, water, and sep tank) Lowest quality: Elec only	0.05

Feature 1 - Overall Quality of the House

Rating of the overall condition of the house from 1 to 10:

- 1 - Very poor condition
- 10 - Excellent condition

Based on our model, **improvement** in the rating by 1 **score** is estimated to increase sale price of **approx. $13,188**

Trend of Sale Price by Overall Quality
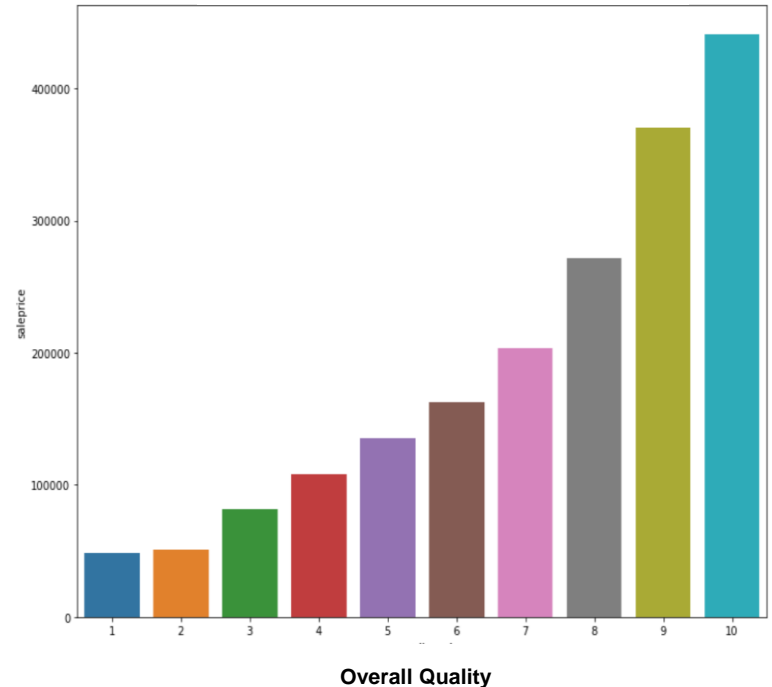
Feature 1 - Overall Quality of the House
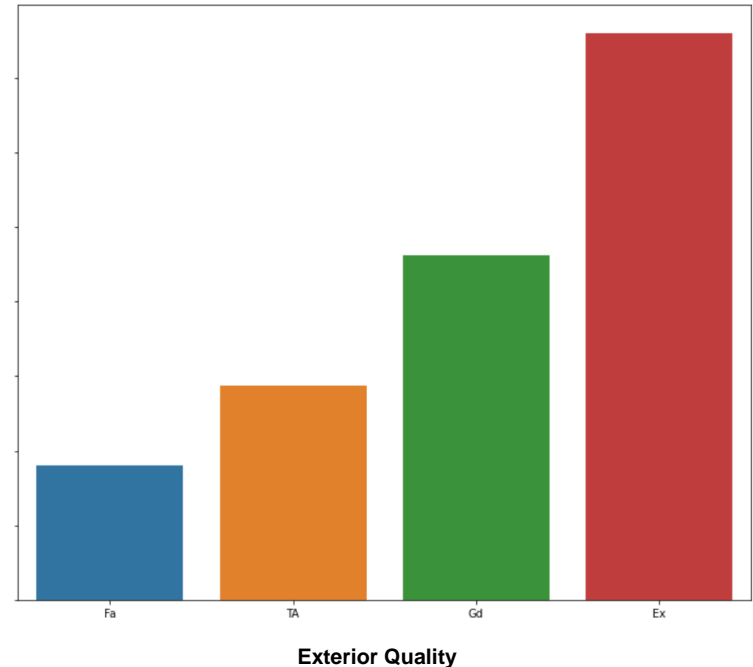

Poor Overall Quality

Excellent Overall Quality

Feature 2 - Exterior Quality

Rating of material quality of the exterior from 1 to 4:

- 1 - Fair
- 4 - Excellent

Based on our model, **improvement** in the rating by 1 **score** is estimated to increase sale price of **approx. $24,450**

Trend of Sale Price by Exterior Quality

Exterior Quality

Feature 2 - Exterior Quality

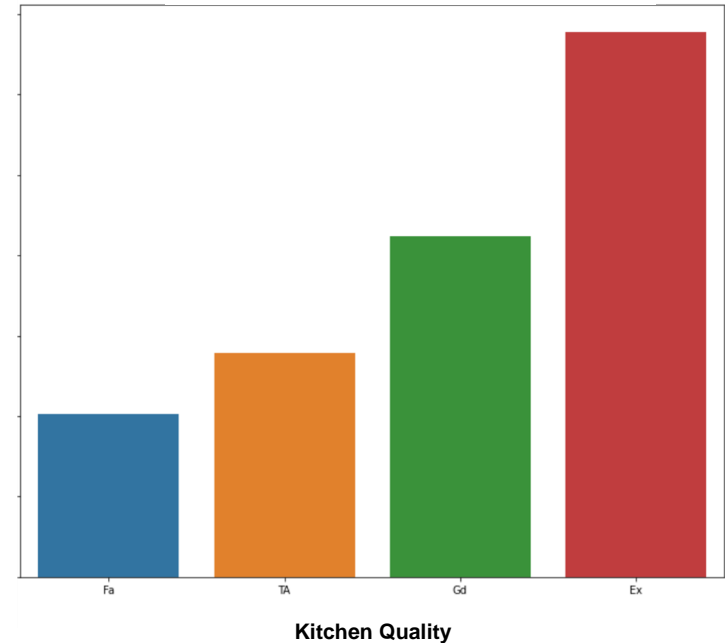**Poor Exterior Quality**

Excellent Exterior Quality

Feature 3 - Kitchen Quality

Rating of the kitchen quality ranges from Fair to Excellent:

- 1 - Fair
- 4 - Excellent

Based on our model, renovating kitchen to an excellent condition is estimated to increase sale price of **approx. $25,000**

Trend of Sale Price by Kitchen Quality

Kitchen Quality

Feature 3 - Kitchen Quality

Poor Quality Kitchen

Excellent Quality Kitchen

Conclusion

3 key renovations we recommend to help property owners maximum housing sale price:

1		The Kitchen	⬆️ ~$25,000 / 1 score increase

2		The Exterior	⬆️ ~$24,000 / 1 score increase

3		The Overall Quality	⬆️ ~$13,000 / 1 score increase

Actionable Insight

Based on the client (house owner) current housing condition, GA marketing team can propose to client, the actual increase in housing price by the recommended types of renovation packages.

Next Steps

Currently, the evaluation of features is manually done by our consultants and the evaluation is based on individual's judgement (subjective) and time consuming.

We will be working on **image processing tool** for our GA consultants **to automate the evaluation of features** for a consistent scoring and reduce the amount of time for the evaluation.

Appendix A - Data Cleaning

S/N	Type	Findings	Action
01	**Missing Value / Imputation** for Numerical Features	● There are **3769** Missing Values! ● However, most of the missingness is expected. It explains the absence of Feature.	● *For categorical/Ordinal, replaces NaN to "Not applicable"* ● *For Continuous/Discrete, replaces NaN to 0* ● Replace missing values with mean (case-by-case basis)
02	**Outliers**	● Large Houses (>5000 square feet) but sold at very low prices ● Data for the year where the garage is built is indicated as year 2207	● *Removed 2 outliers which are large houses sold at very low prices* ● *Replaced 1 future-dated year the garage is built to the year the house is built*
03	**One hot encoding** for Categorical Features (e.g. quality of basement, fireplace) -	● Assume missing values as not having the features	● *Fill Missing values with None*

Appendix B - Additional Considerations

Other features which can also improve housing price:

- Size of Garage - Area, Car Capacity
- Basement Height and Size

Violin plot of Sale Price vs Number of Cars in Garage

Scatterplot of Tot_bsmt_sf vs Sale Price