

# Our Agent of Change will Move You

*Real Estate Consultants :*

*Steven Lee | Zawanah Sainuddin | Adrian Teng*

# Agenda

1. Objectives
  - a. Business Objective and Problem statement
2. Insights
  - a. Polynomial Features
3. Model
  - a. Model Evaluation
  - b. Benchmark
4. Recommendations

# Problem Statement

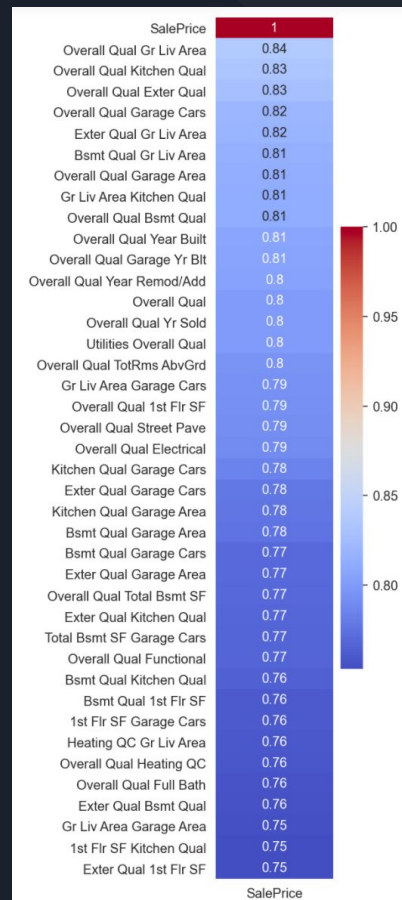
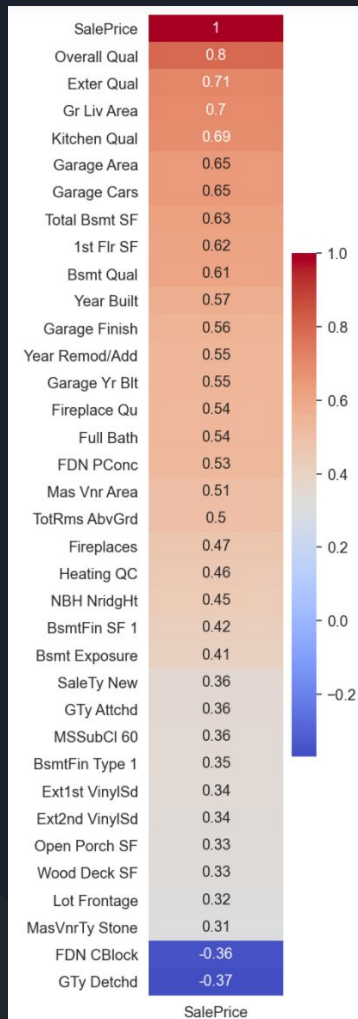
- We have been tasked to develop a regression model to predict sale prices of residential properties based on housing data from the city of Ames in Iowa
- Determine the three most important features that drive house prices.
- Primary stakeholders : Consulting firms
- Secondary stakeholders : House owners looking to sell their houses (i.e. consulting firm clients)



# Polynomial Features AKA Interaction Terms

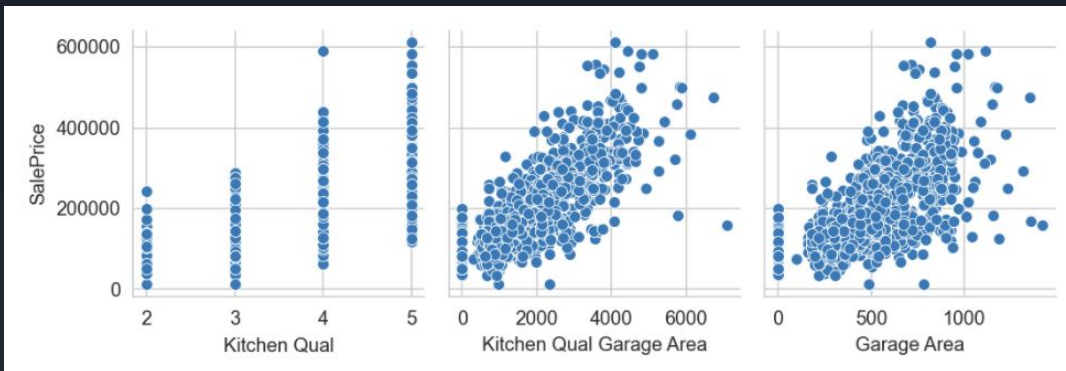
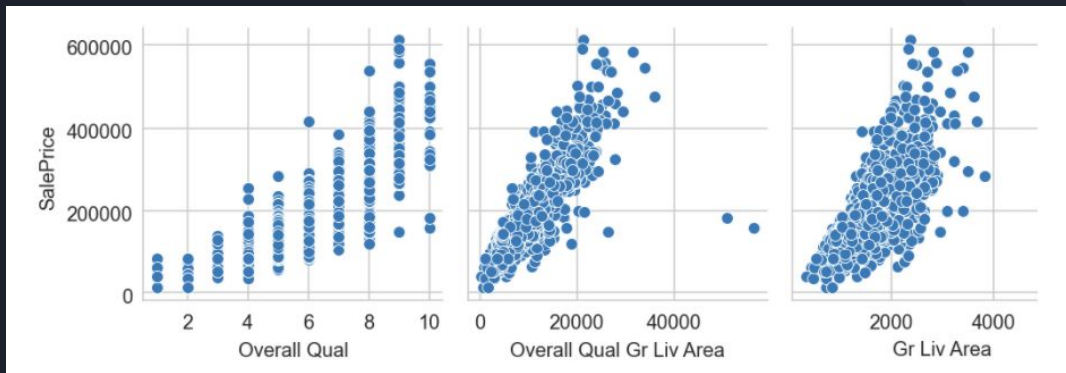
What? Why?

# Correlation Comparison before & after Polynomial Features



# Interaction Term Examples

- Overall Qual Gr Liv Area
- Kitchen Qual Garage Area



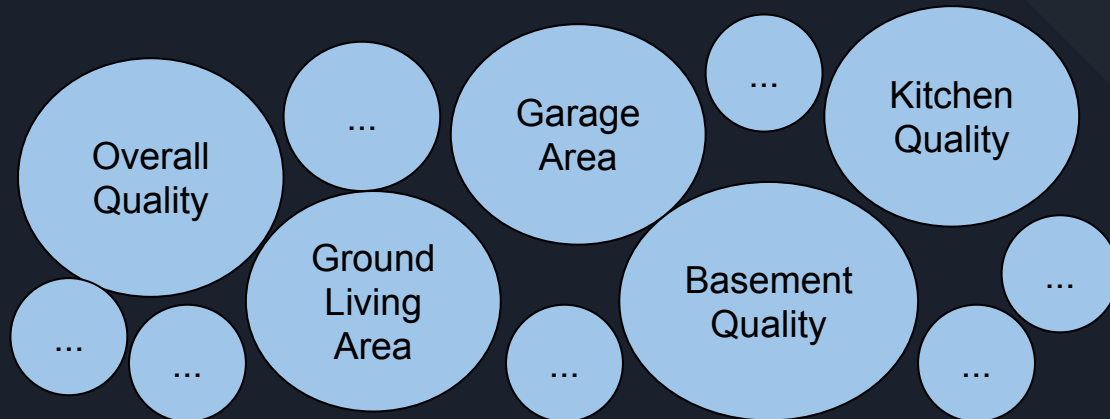
*Agenda*

*Insights*

***Model***

*Recommendations*

# Modelling



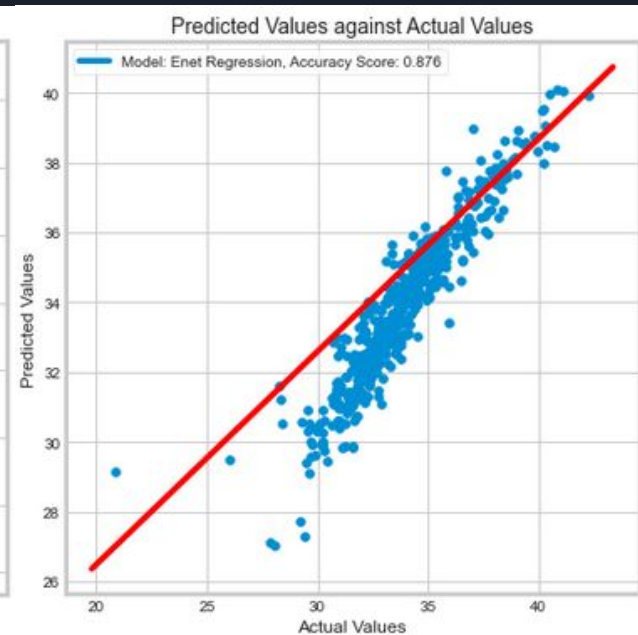
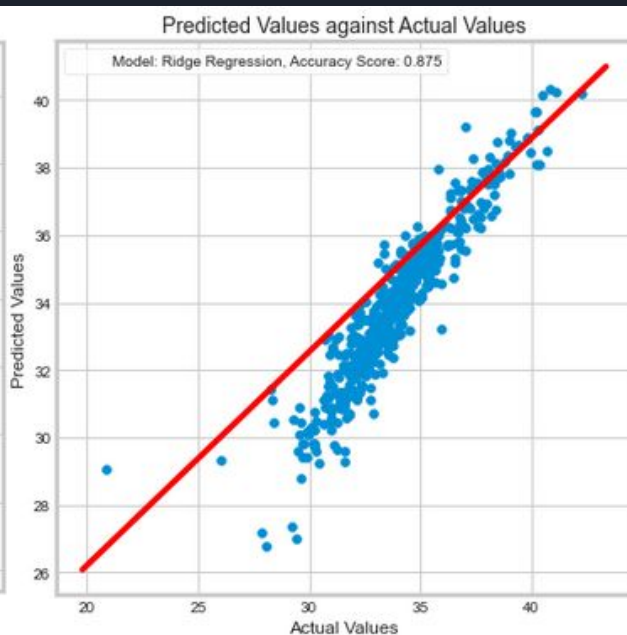
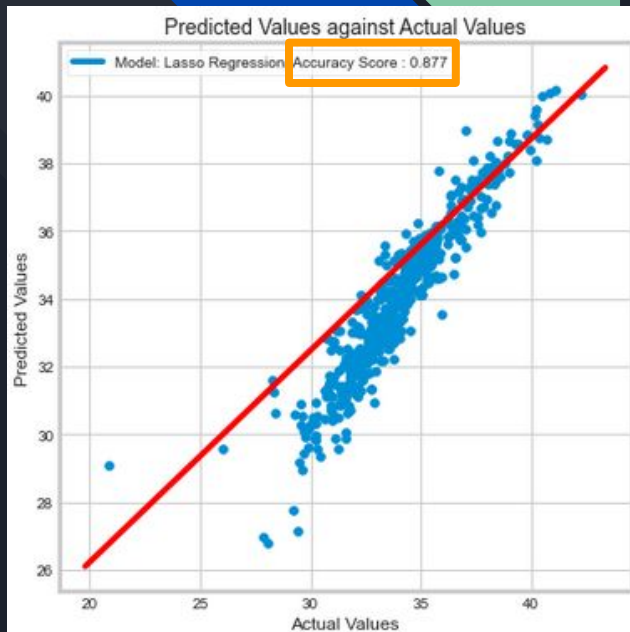
- Fit all features into models
- Baseline model as benchmark for how good the production model is

## **Models**

- Lasso Regression Model
- Ridge Regression Model
- Elastic Net Model
- Baseline Model

# Model Selection

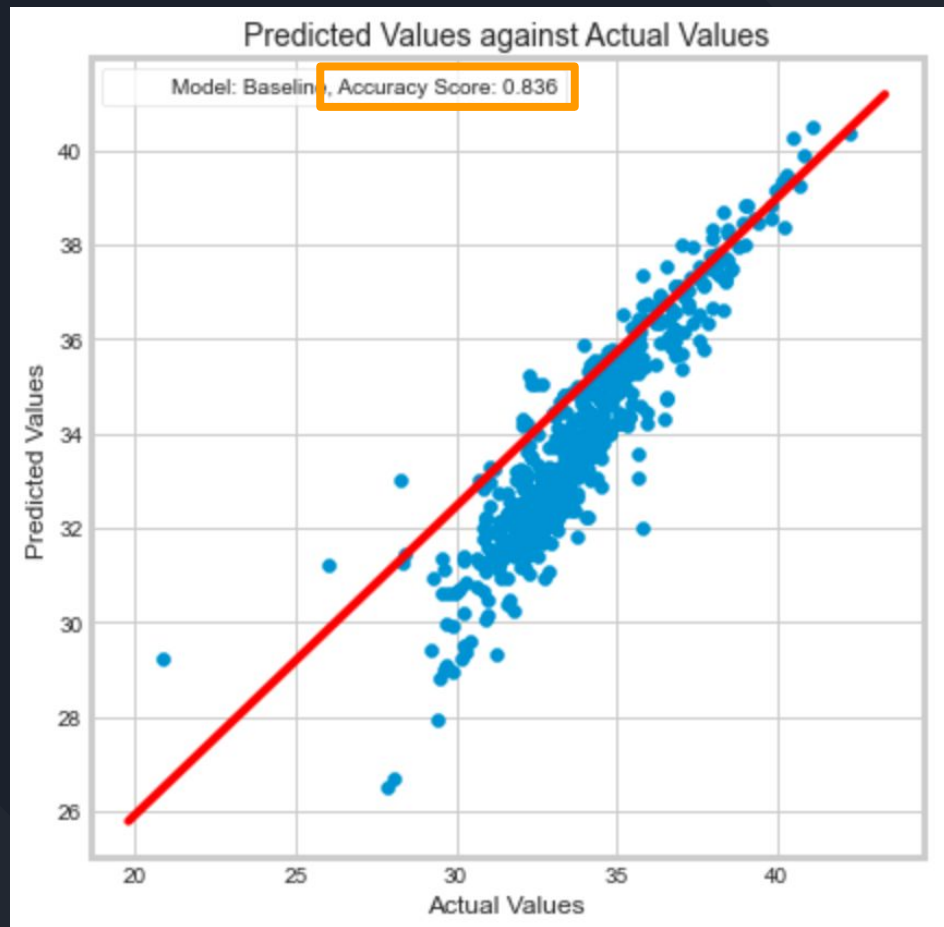
- Success metric : R2 score (ie. Accuracy Score)
- Accuracy Score closer to 1 represents a better fit
- Lasso is the best performing model among others, being able to explain **87.7%** of variability in data.





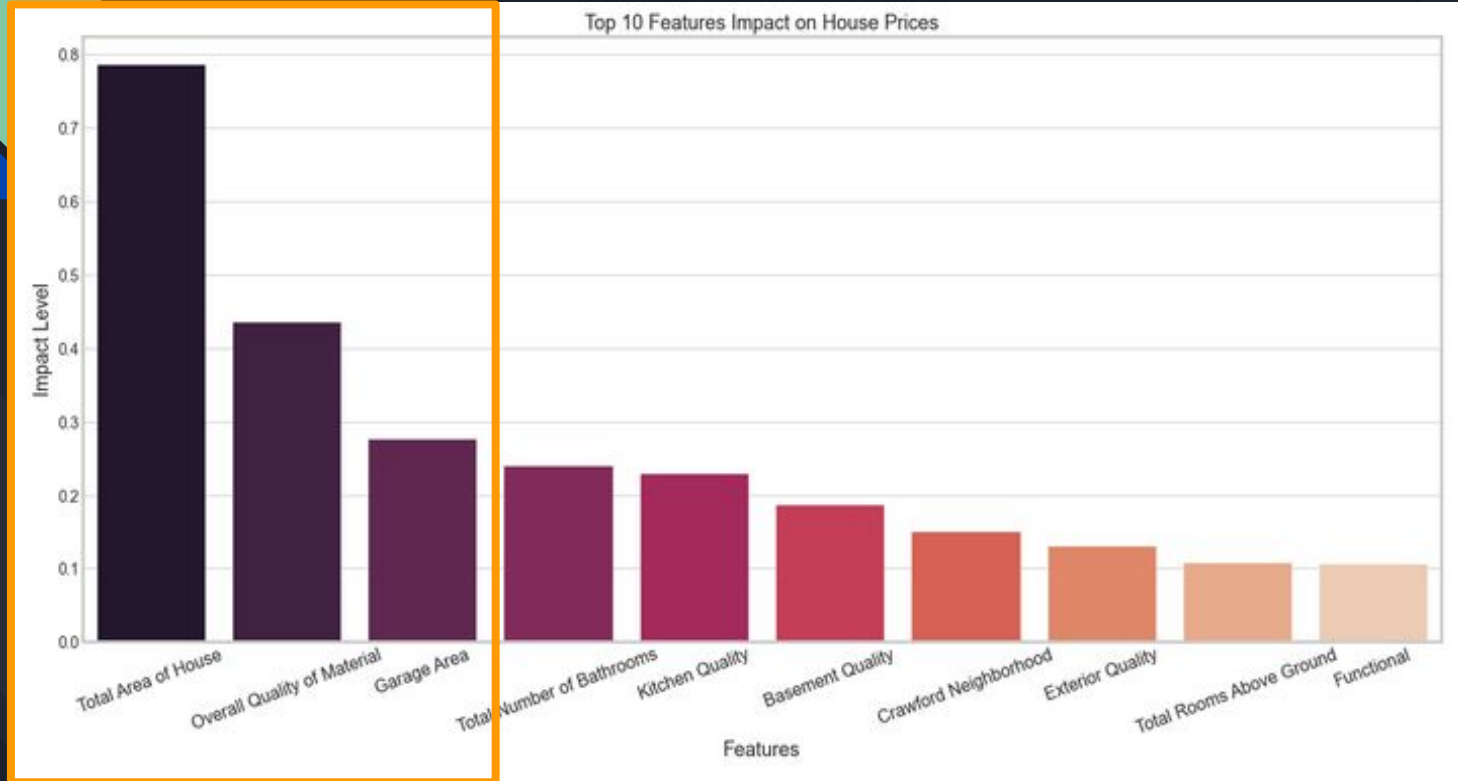
# Baseline Model

- Baseline model : Lasso with top 5 numeric features with strong correlation with sale price
- Accuracy score of 0.836 (ie. able to explain 83.6% of variability in data)
- Lasso Model performed better than benchmark and will be our production model



# Top 3 Features

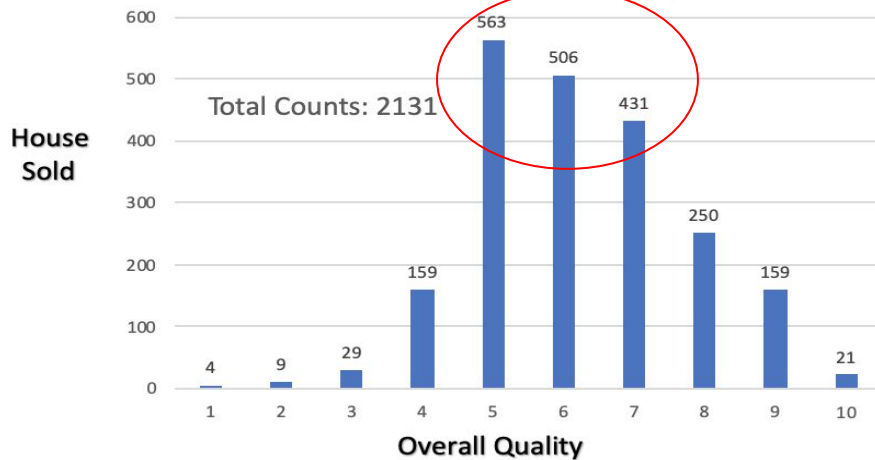
- Top 3 features that drive house prices are total area (ground living area and basement area), overall quality and garage area.



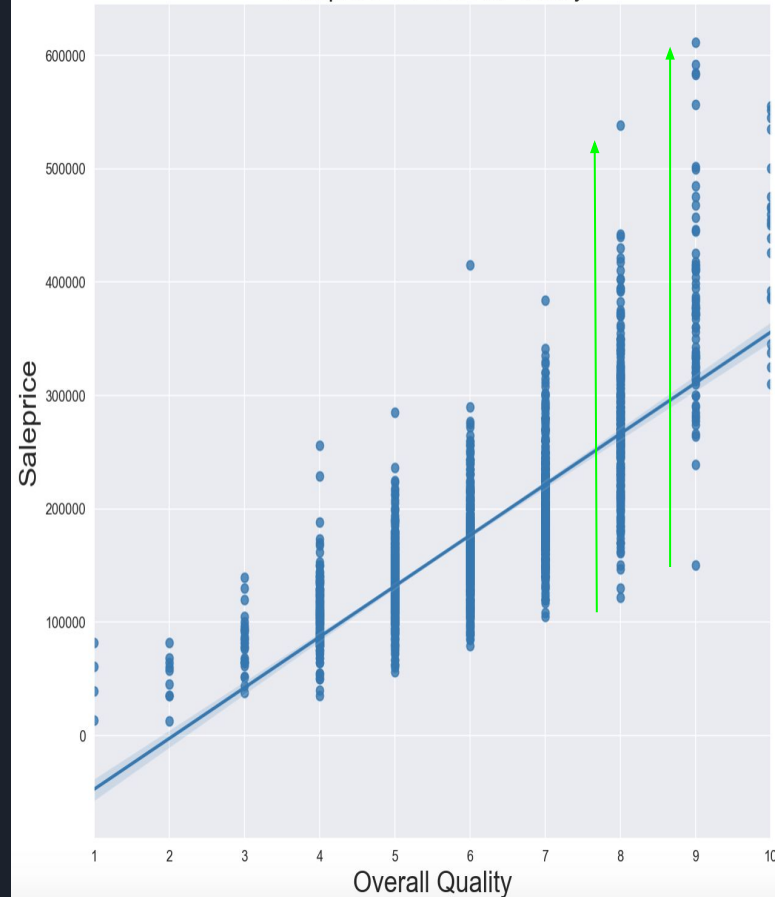
# Overall Qualities

- Most of the house sold is at 5-7, at the price range between 50k-400k.
- 8-9, have the highest spread of price

House Sold for difference level of qualities

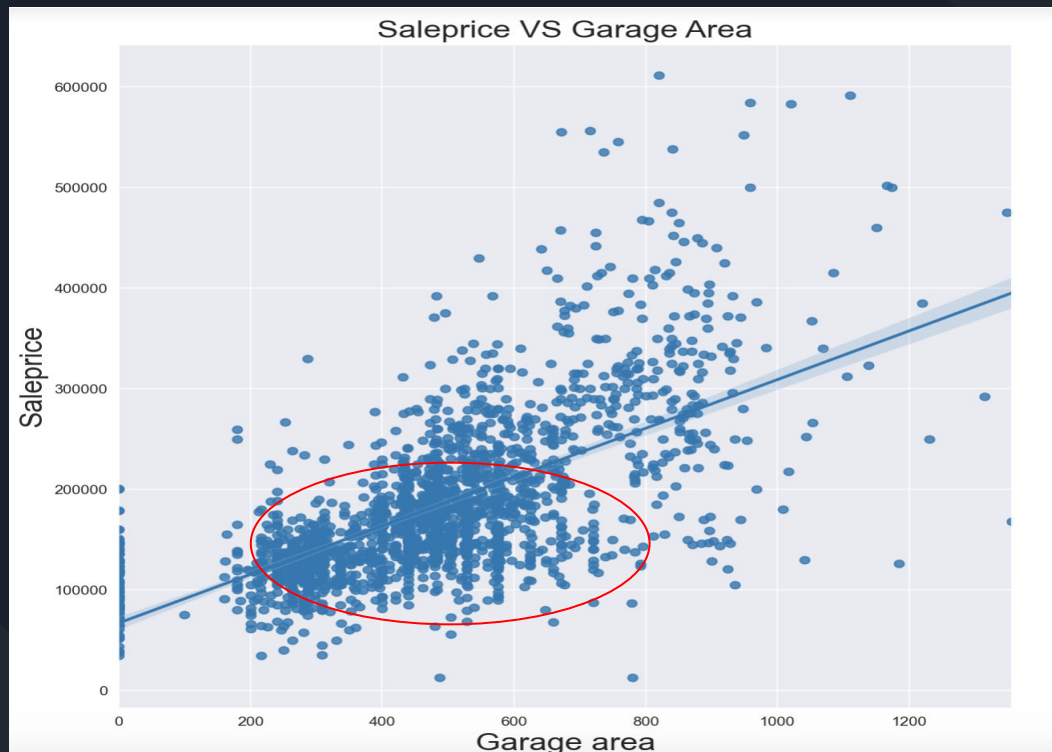
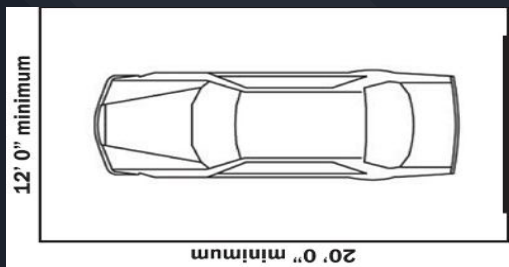


Saleprice VS Overall Quality



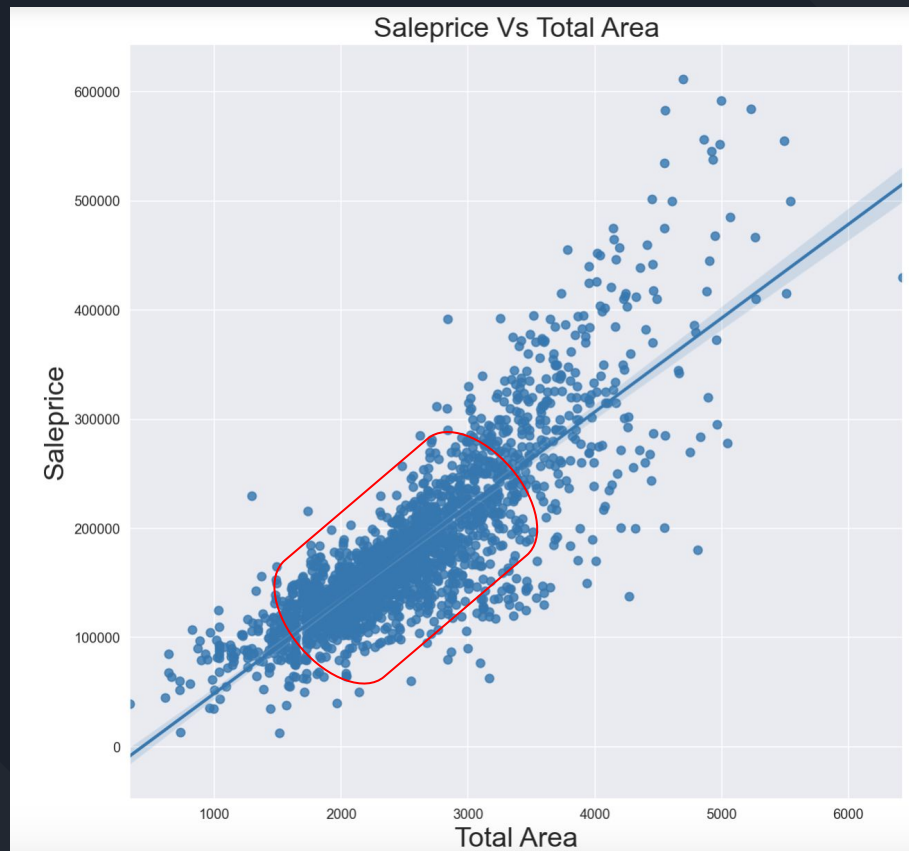
# Garage Area

- High concentration around the area between 200-600 sq ft, it is also around 80k-260k



# Total Area

- Very high correlation
- High concentration 1800 to 3600 sq feet, around 100k to 360k



# Recommendations

Features	Correlation with sale price (1 is highest)	Advice to Home Owners who wish to sell	How to Sell to Home Buyers	Recommendations according to demand
Overall Qualities	2	Ensure that fixtures are working well. No spoilt items, before viewing by potential new owners.	Emphasise quality of furnishings or area, at least in average condition.	Keeping the quality average to good conditions. Not too excessive to the price.
Garage Area	3	If too huge, consider changing garage space to a living area, eg. expand basement size.	Emphasize fit of garage area for number of cars for the particular buyer	200-600 sq feet, good to fit 1-2 cars for an ordinary family.
Total Area	1	Ultimately, the bigger the area the higher the value.	Facilities required and financial abilities.	Separate the land to 2 units according to 1800 to 3600 sq feet or adding essential facilities, if there is excessive space