2010 Ames Property Market Sharing - By 5AM Club

PROJECT 2 - Ames Housing and Kaggle Challenge

By 5AM Club
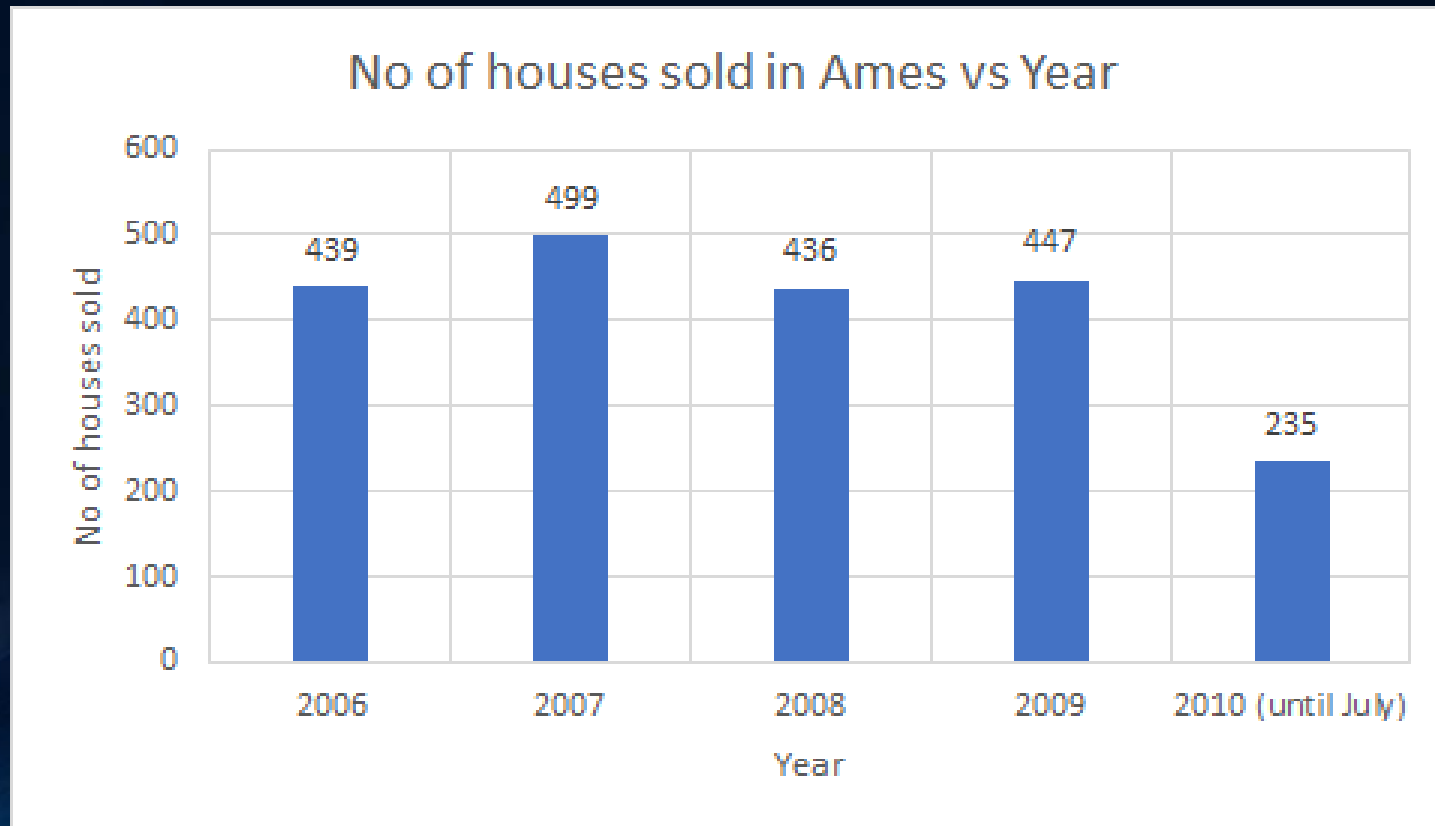1) Andrew Chia
2) Deepankar Sharma
3) Tan Wee Jin
4) Tay Yi Li

Background

- 5AM Club is a Data Science Agency that specialises in creating customised business solutions for its clients.

- Some of our clients include the United States Federal Government, the United States Men's National Soccer Team and even real estate tycoon Donald Trump (we were not fired!)

- Today, based on data from houses sold in Ames in 2006-2010, we are going to share some important insights from our study to help you, the realtors of Skywalker Property Advisors, reach greater heights!

Background

- Situation of the property market in Ames:

  - Sales have been relatively stable across the last 4 years, despite experiencing a subprime Financial Crisis since 2008.

No of houses sold in Ames vs Year

Year	No of houses sold
2006	439
2007	499
2008	436
2009	447
2010 (until July)	235

Content

- Problem Statement

- Methodology

- Findings and Recommendations

- Conclusion and Future Work

Problem Statement

"This project aims to identify areas contributing to high transacted prices and where the highest transacted volume occurs, using a data science approach, so as to help realtors of Skywalker Property Advisors gain a competitive advantage in the Ames Housing Market."

Methodology

Dataset we have used for analysis:
- 2051 transaction entries describing the sale of individual residential property in Ames, Iowa from 2006 to July 2010.

Done extensive data cleaning and preprocessing prior to modelling, such as imputation of missing values, feature engineering, and scaling.

Then, we use 3 different regression models
- Linear regression
- Ridge regression
- Lasso regression

for modelling of data, before arriving at the best model for our findings.

Methodology

Apart from statistical modelling techniques, the team also considered factors that are anecdotally important to prospective home-buyers. Some of these include neighbourhood, house quality, and house amenities. We will be exploring these in the coming slides.

There are some factors which are perceived to be important by realtors, however, the data tells us otherwise. We will also be looking at factors which have a negligible effect on sale price.

Finally, we will look at the type of houses that are most commonly sold across various categories.

Findings and Recommendations

Using ridge model, we have obtained the top 10 factors (coefficients) which impact the housing sale price in a positive and negative manner.

Top 10 Positive and Top 10 Negative Coefficients
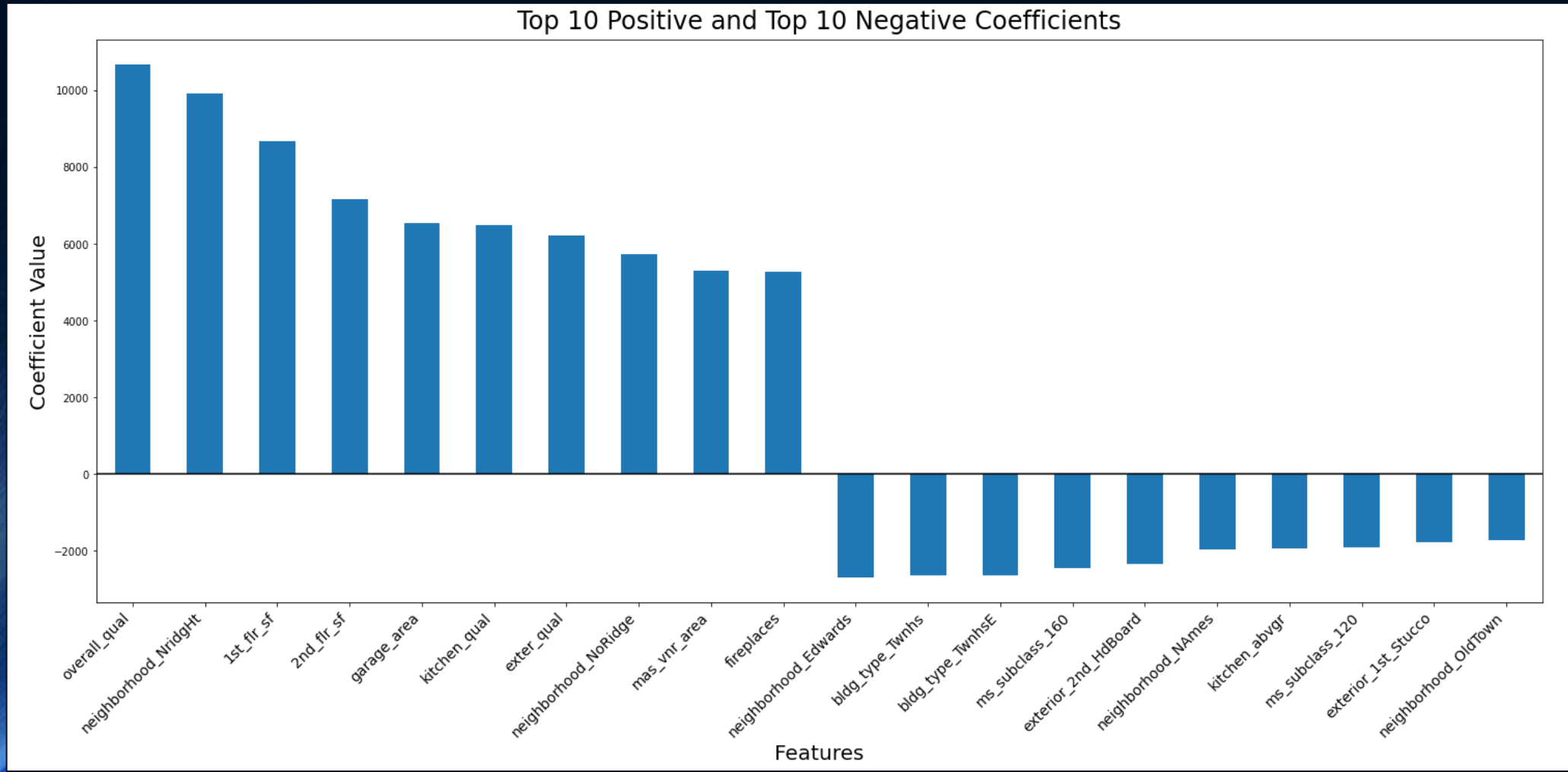
Neighborhood

2 neighborhoods that have +ve impact: Northridge Height and Northridge.
3 neighborhoods that have -ve impact: Edwards, North Ames and Old Town.
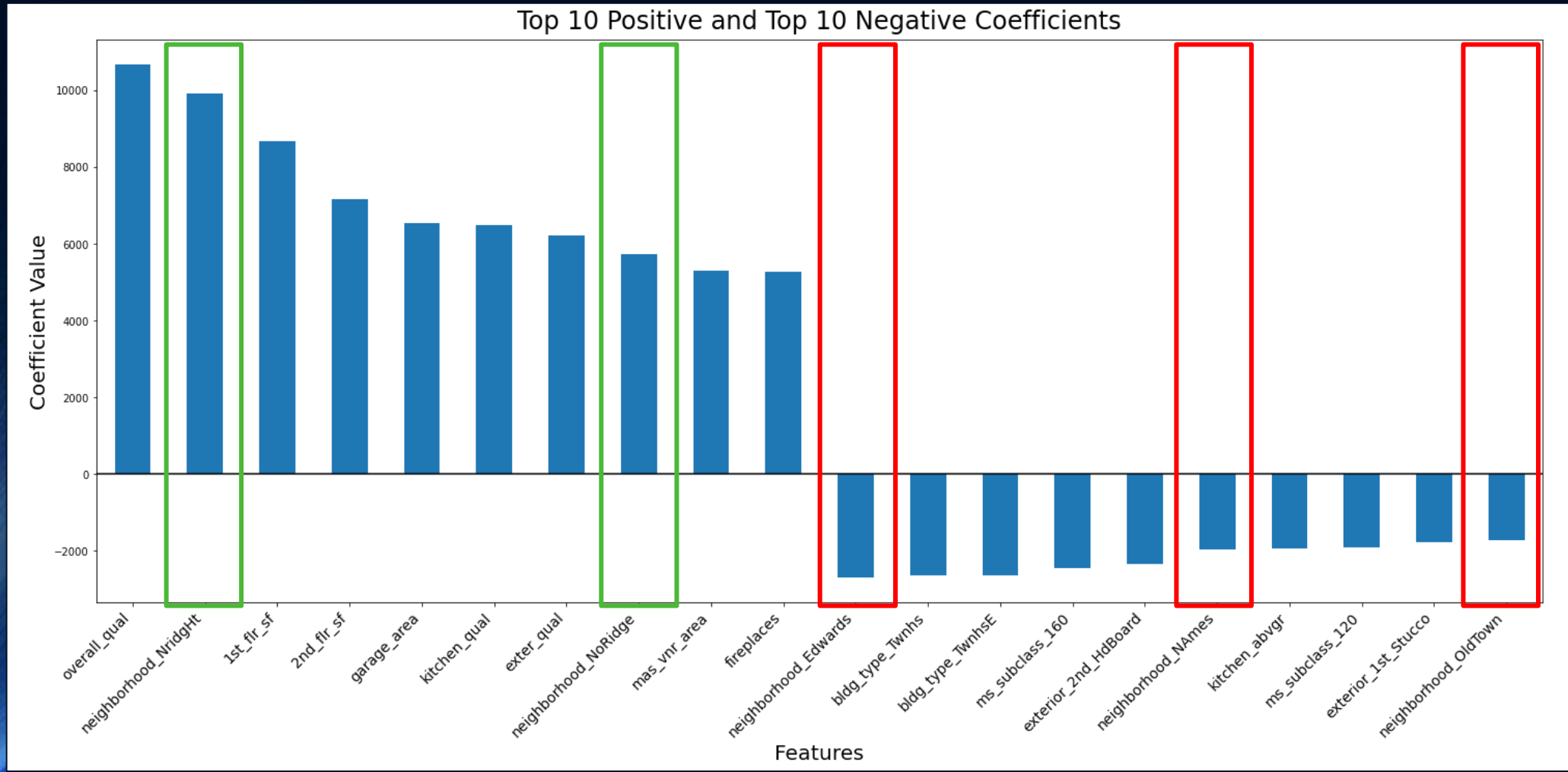
Neighborhood

Observations based on Saleprice
- Houses in Northridge Heights and Northridge have mean sale price of >$300,000, yielding higher amount of revenue and commission.
- Houses in North Ames, Edwards, and Old Town have mean sale price of <$150,000, yielding lesser amount of revenue and commission.
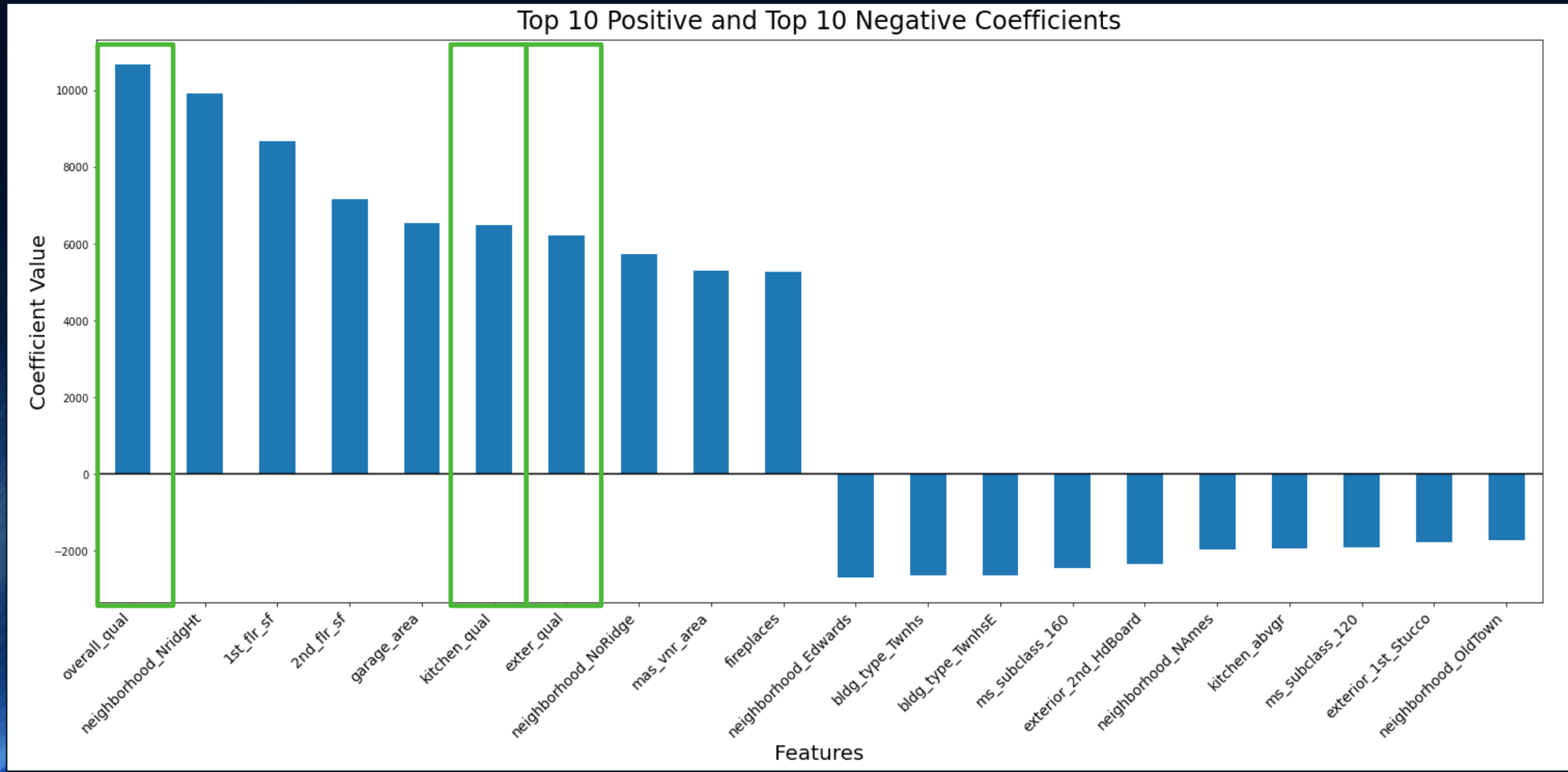
Recommendation
- Realtors could focus more on marketing the houses in Northridge Heights and Northridge to generate higher revenue through higher sale prices.
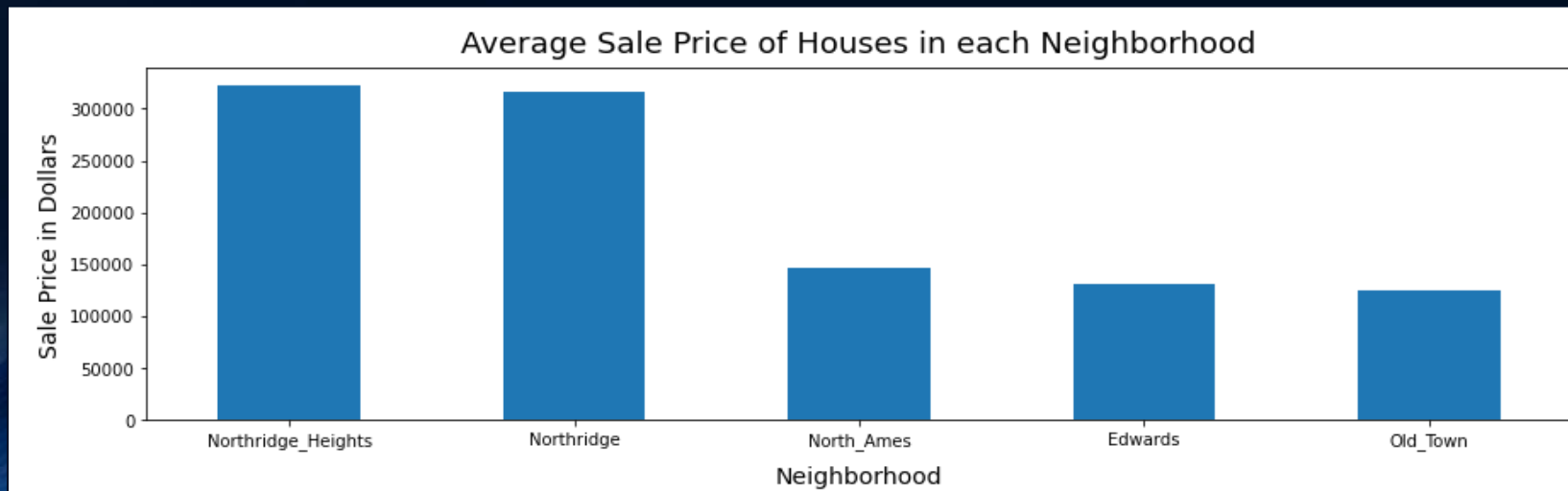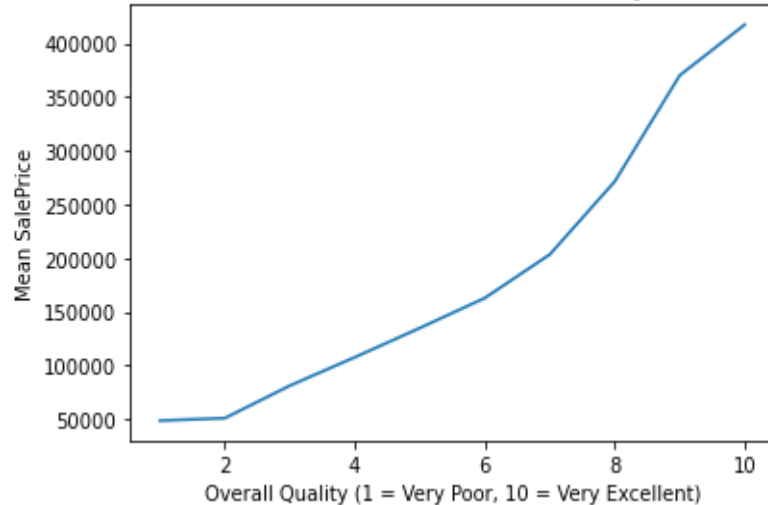
Quality of House

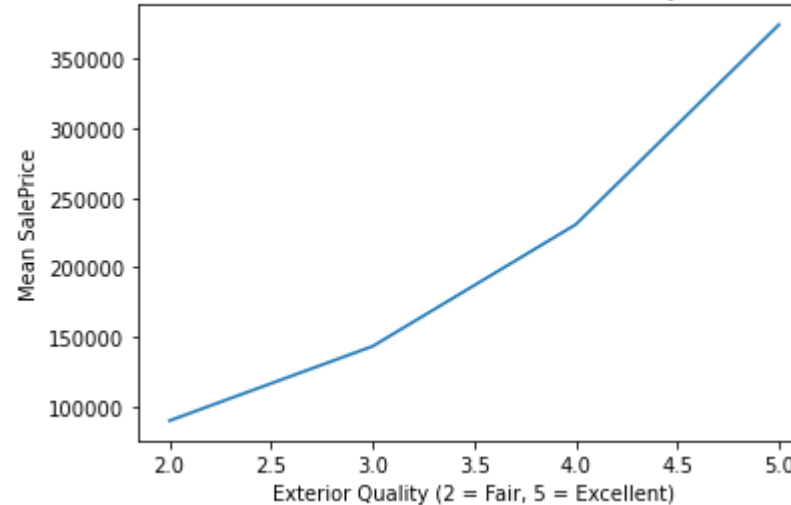The Overall, Exterior and Kitchen Qualities of the house all have a positive impact on sale prices.

Quality of House

The higher the Overall, Exterior, and Kitchen Qualities of the house, the higher the mean sale price of the house.
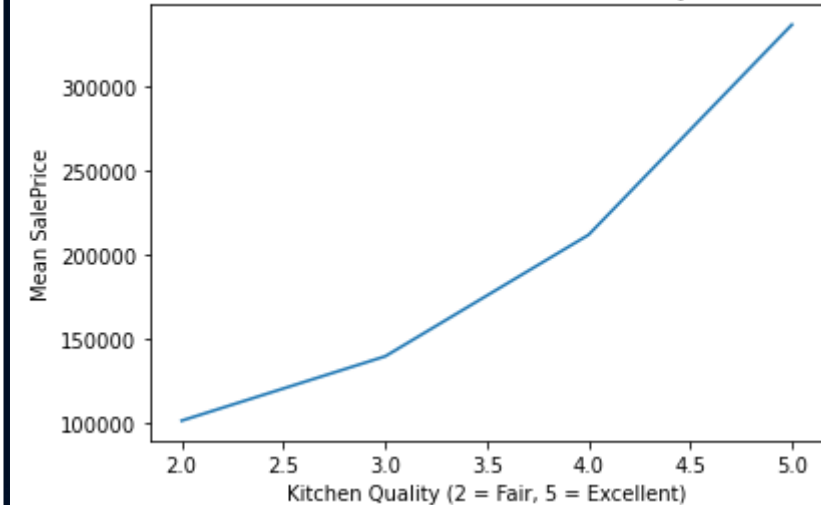
Quality of House

Current situation: Most number of houses sold -
have 3 - 4 (Average - Good Rating) for exterior quality and kitchen quality,
and a range of 5 to 8 (Average- Very Good Rating) for overall quality.

Recommendation
- For clients with below average housing quality, advice them to refurbish their home
- For clients with average/good housing quality, can try persuading them to improve housing quality even more to differentiate themselves for a higher sale price.
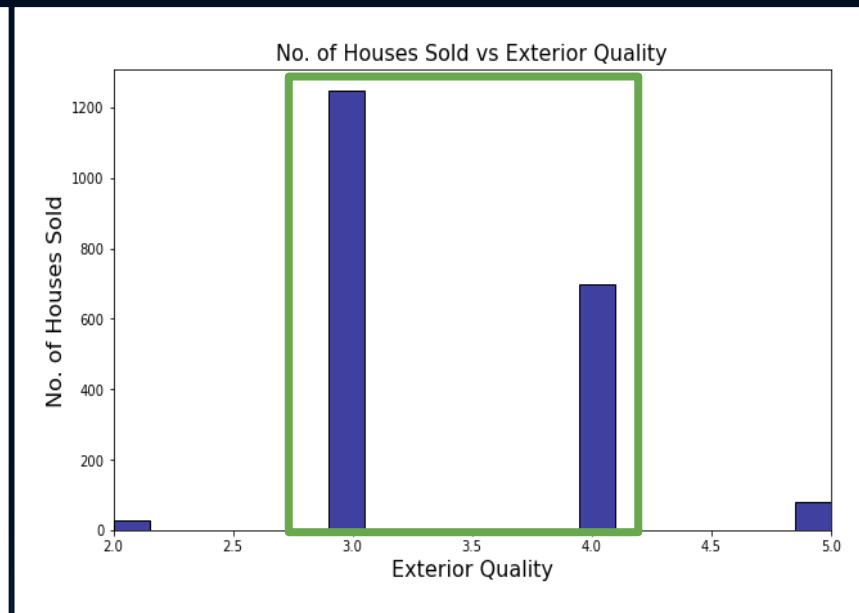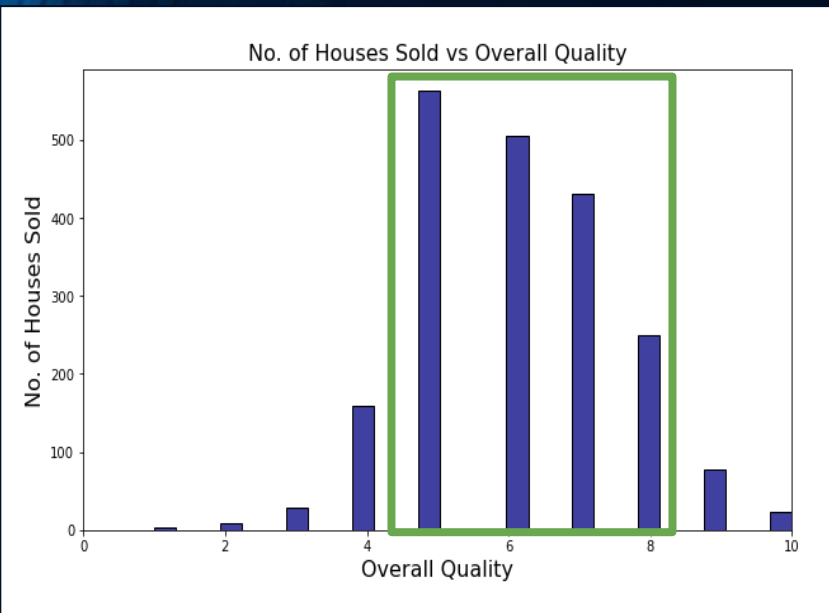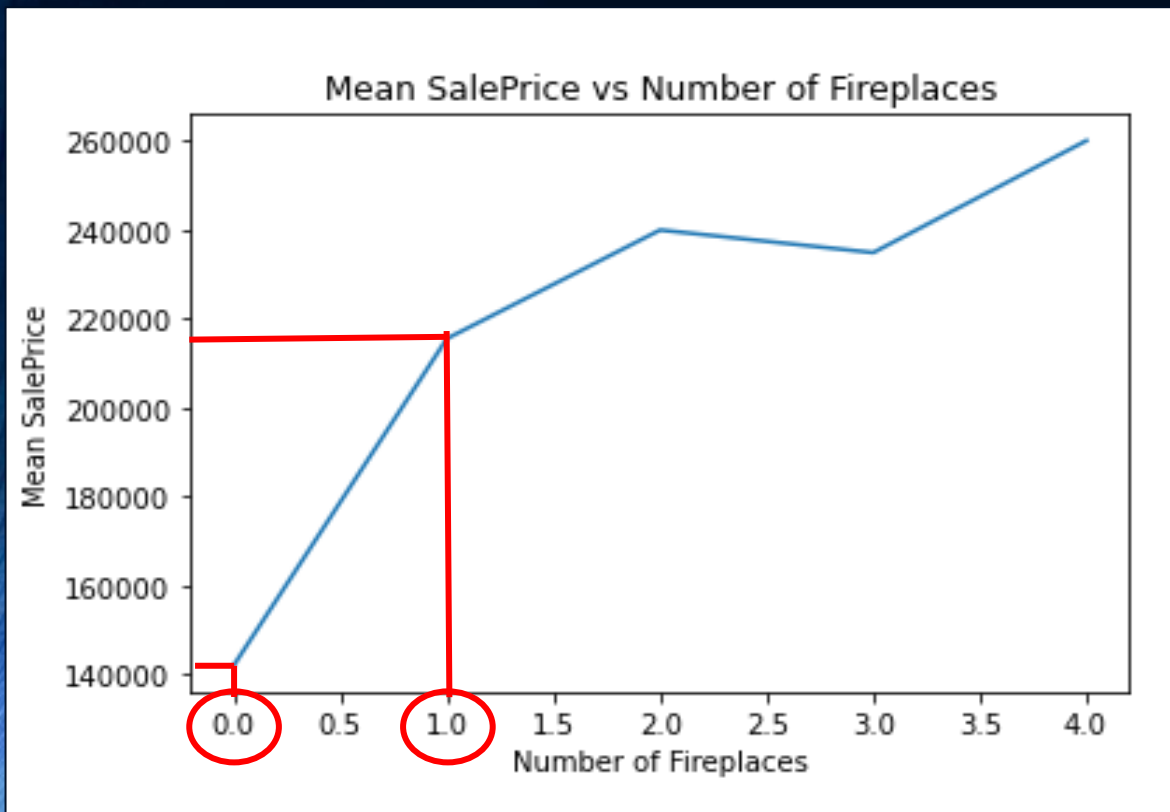
No. of Fireplaces increases Sale Price

Top 10 Positive and Top 10 Negative Coefficients
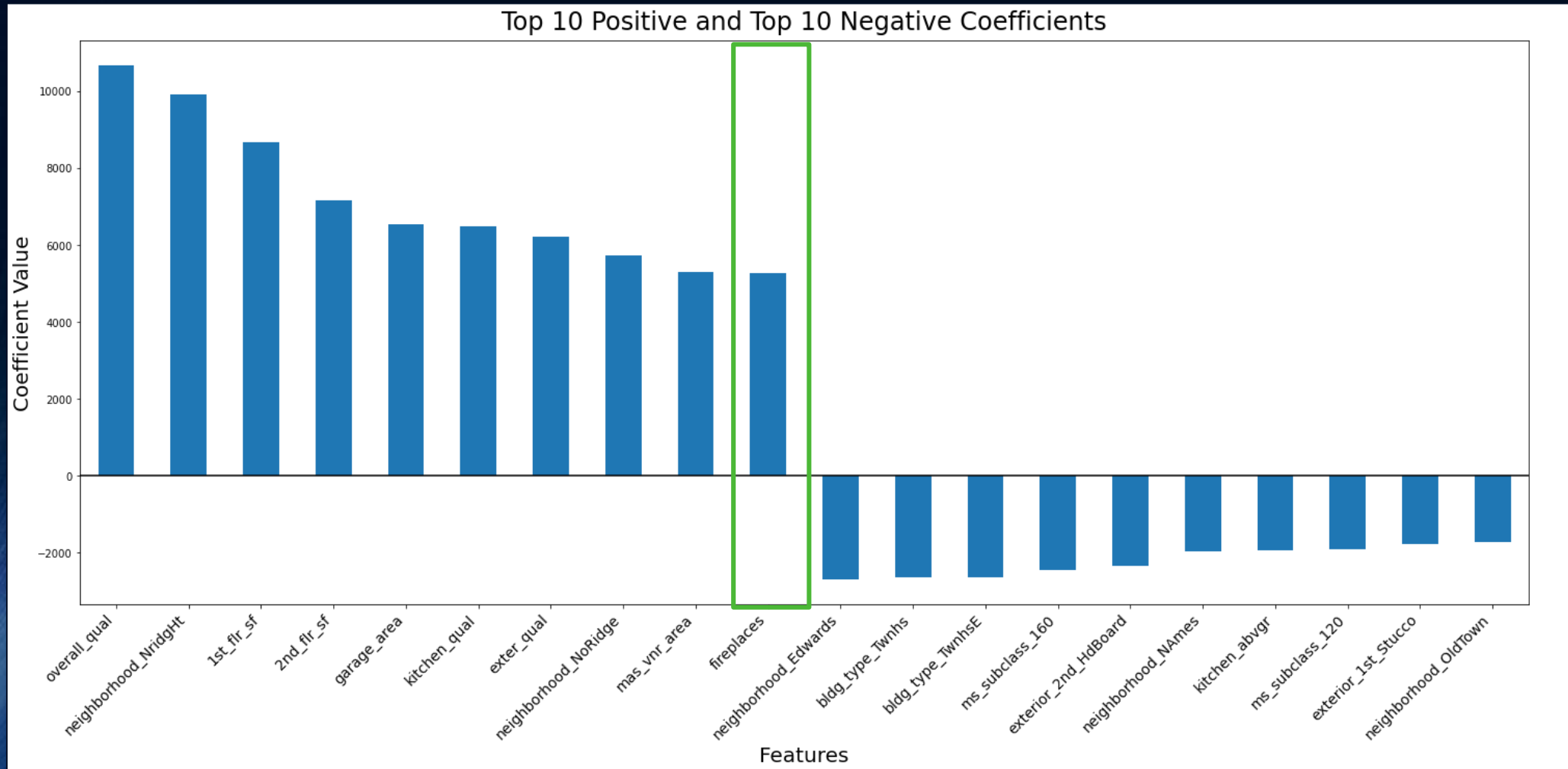
Fireplaces are important in Ames, Iowa

Largest jump in sale price from 0 to 1 fireplace

Mean SalePrice vs Number of Fireplaces

Climate Averages	Ames, Iowa	United States
Rainfall	35.7 in.	38.1 in.
Snowfall	29.6 in.	27.8 in.
Precipitation	108.7 days	106.2 days
Sunny	202 days	205 days
Avg. July High	84.1°	85.8°
Avg. Jan. Low	10.7°	21.7°
Comfort Index (higher=better)	6.9	7

Source: Ames, Iowa Weather against US average

Features to Avoid



Top 10 Positive and Top 10 Negative Coefficients
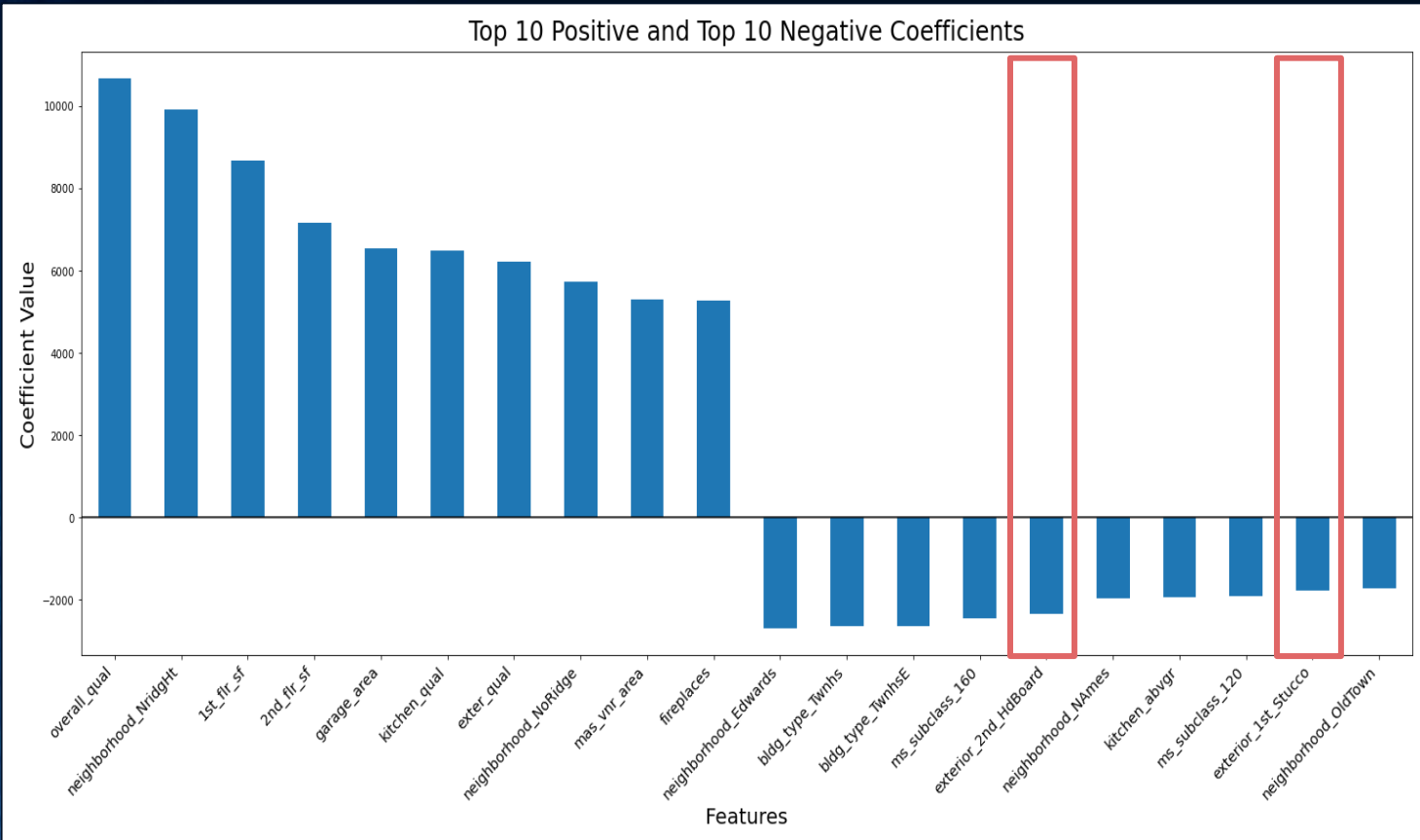
Exterior 2nd HdBoard:
- Poor moisture resisting properties
- Requires periodic inspection and maintenance

Exterior 1st Stucco:
- Wooden frame underneath
- Does not do well in wet climates (eg. high snowfall)

sources: Stucco, Hardboard

Lot Shape and Fence do not affect Sale Price

Features with Absolute Coefficients lower than 100
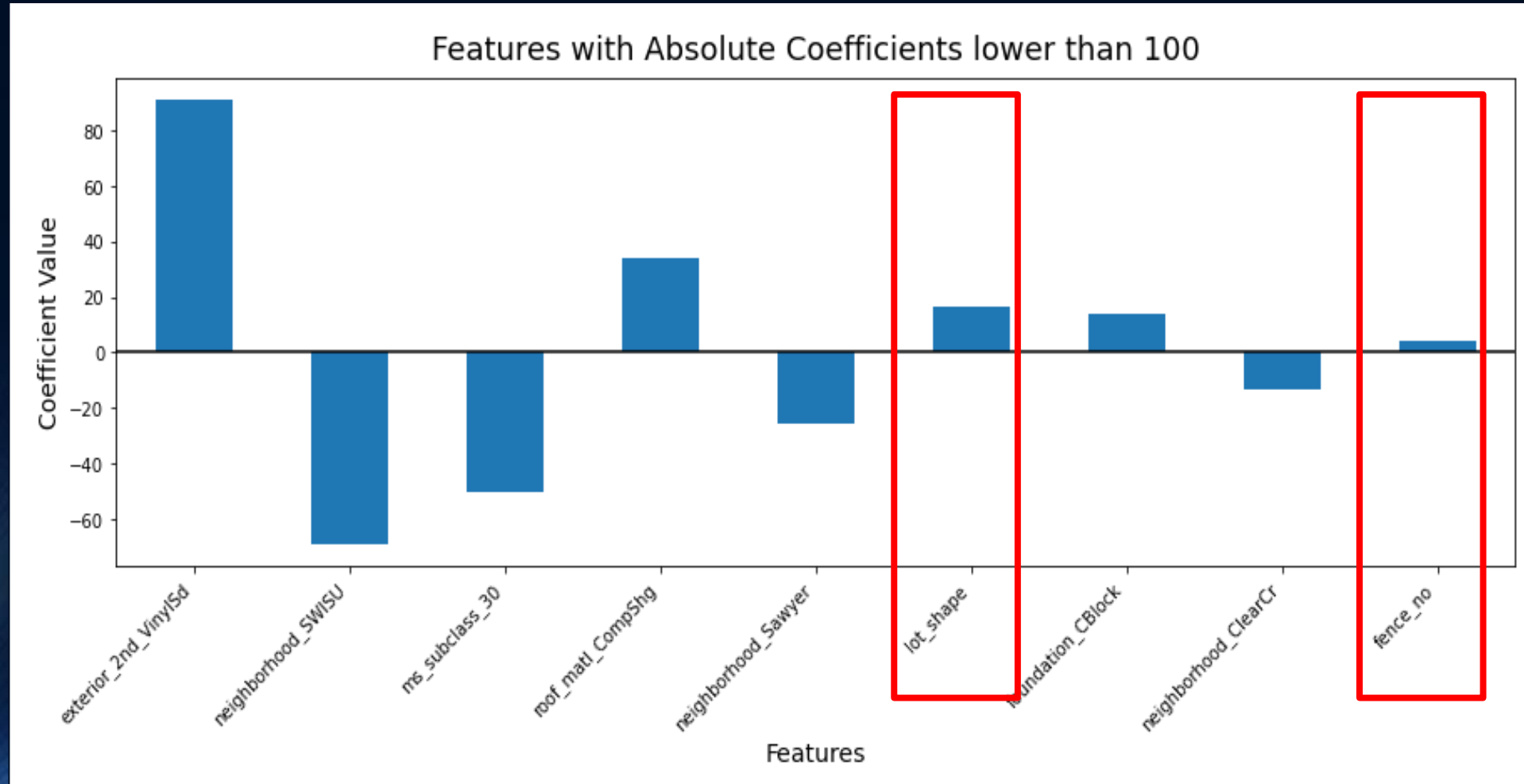
Revenue Generation

There are **TWO WAYS** realtors can increase their earnings:

1. Sell houses that have a higher sale price evaluation

1. Sell MORE houses

   - What are the features in a house that have led to higher number of houses being sold over the past 4 years?

   - Is BIGGER always BETTER?

Square Footage affects Sale of House

| 1st floor | 2nd floor | Garage |

Top 10 Positive and Top 10 Negative Coefficients
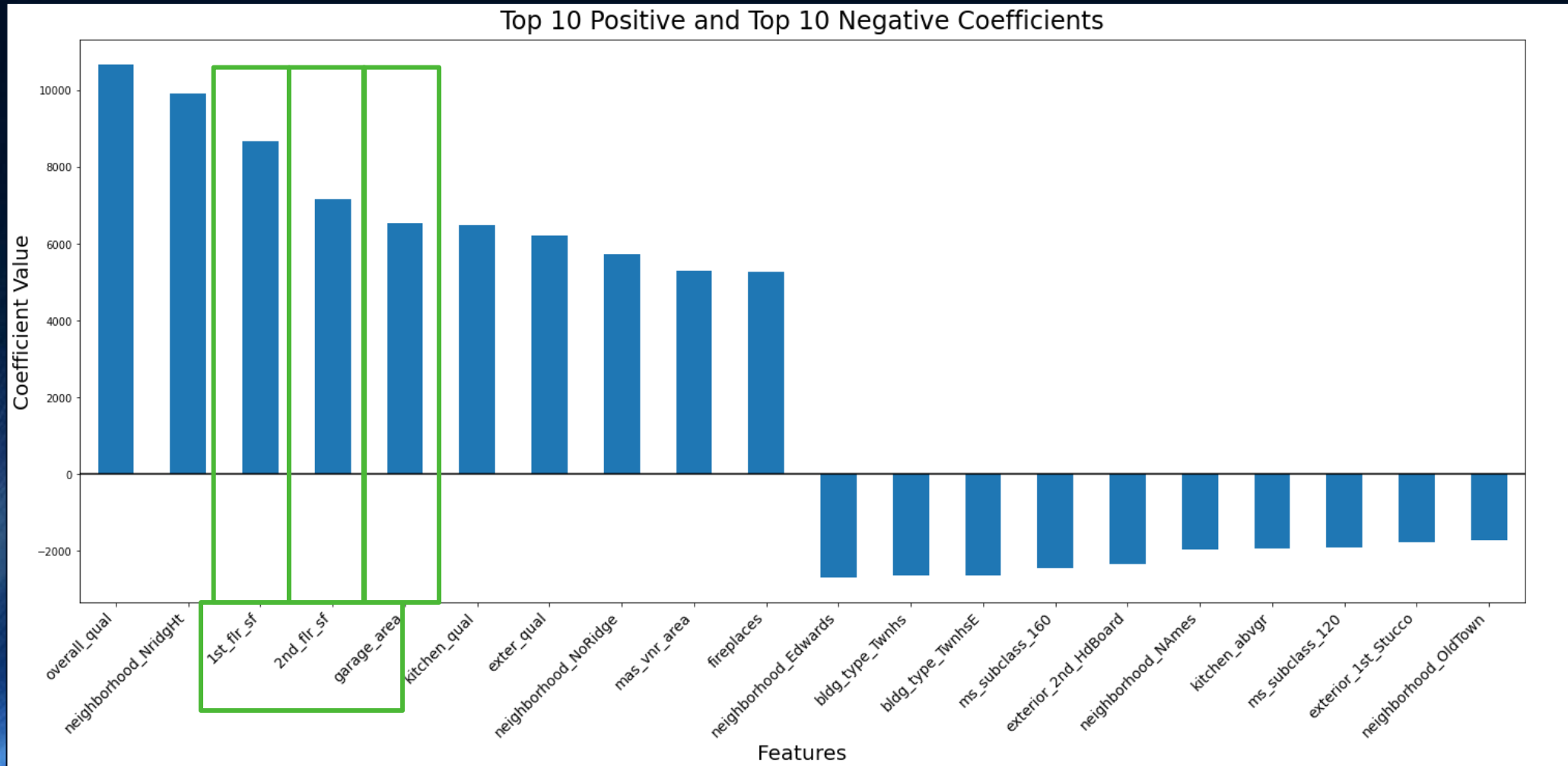
1st Floor Square Footage

- A larger 1st Floor Square Footage generally fetches a higher sale price.

- Houses with a 1st Floor approx <u>800 - 1200 sqft</u> have been the most commonly sold, with houses **800 - 900 sqft** being the standout performer.
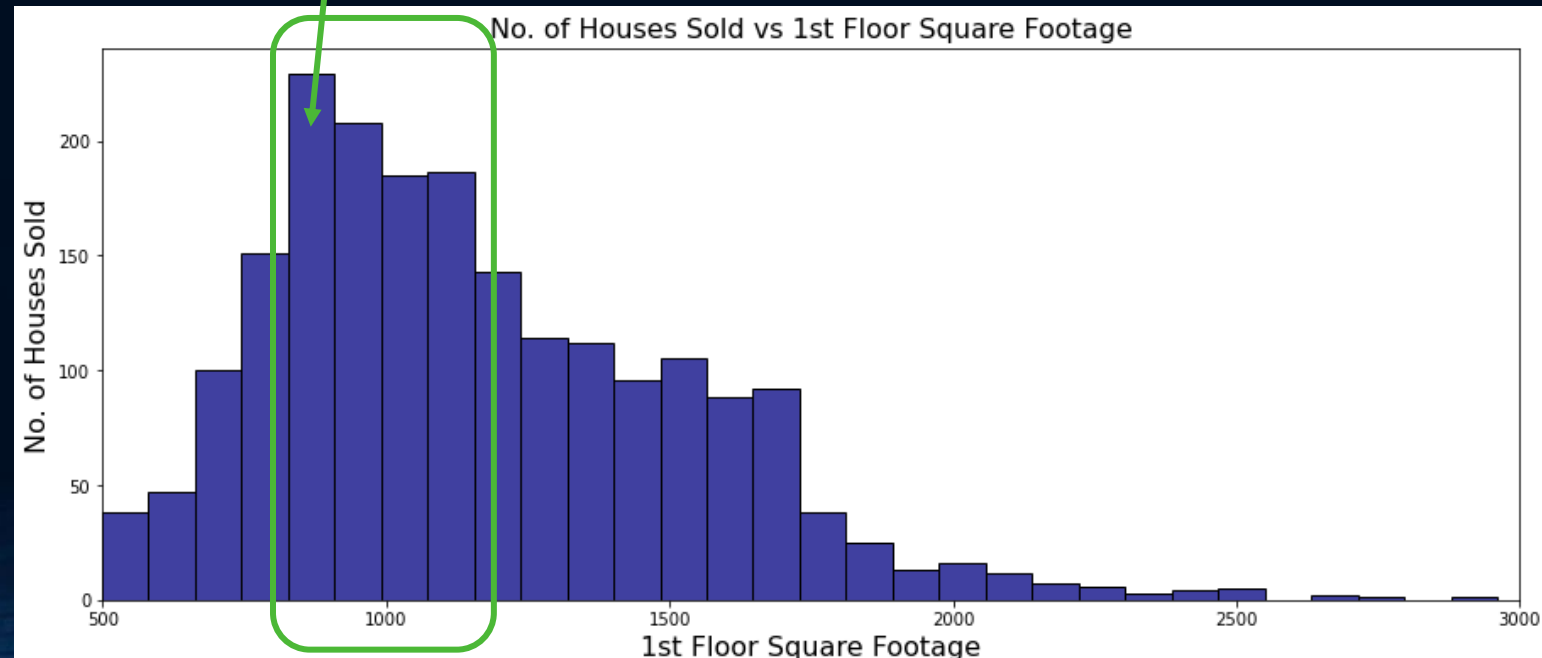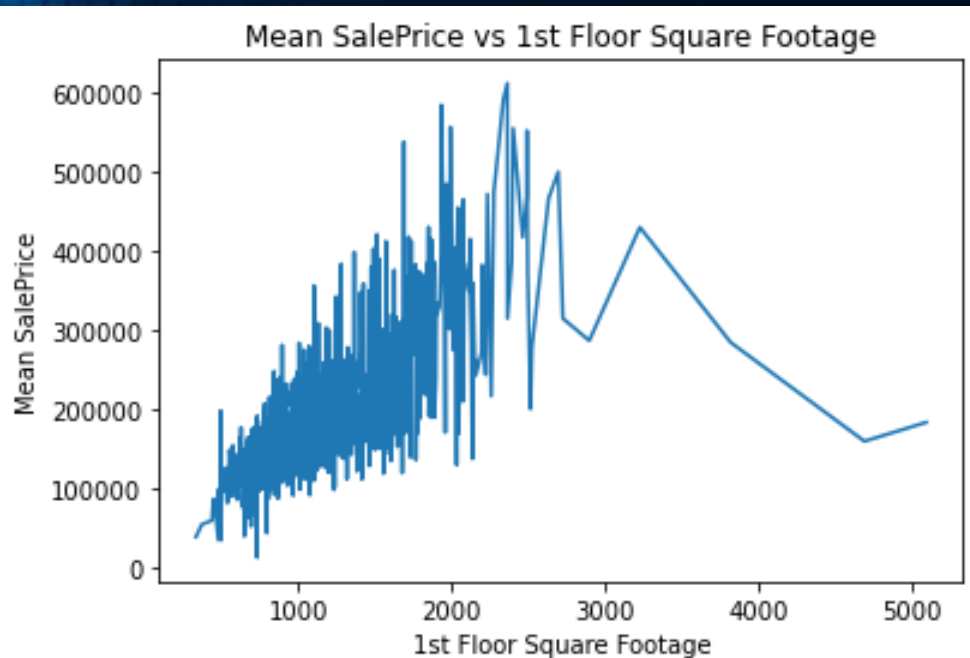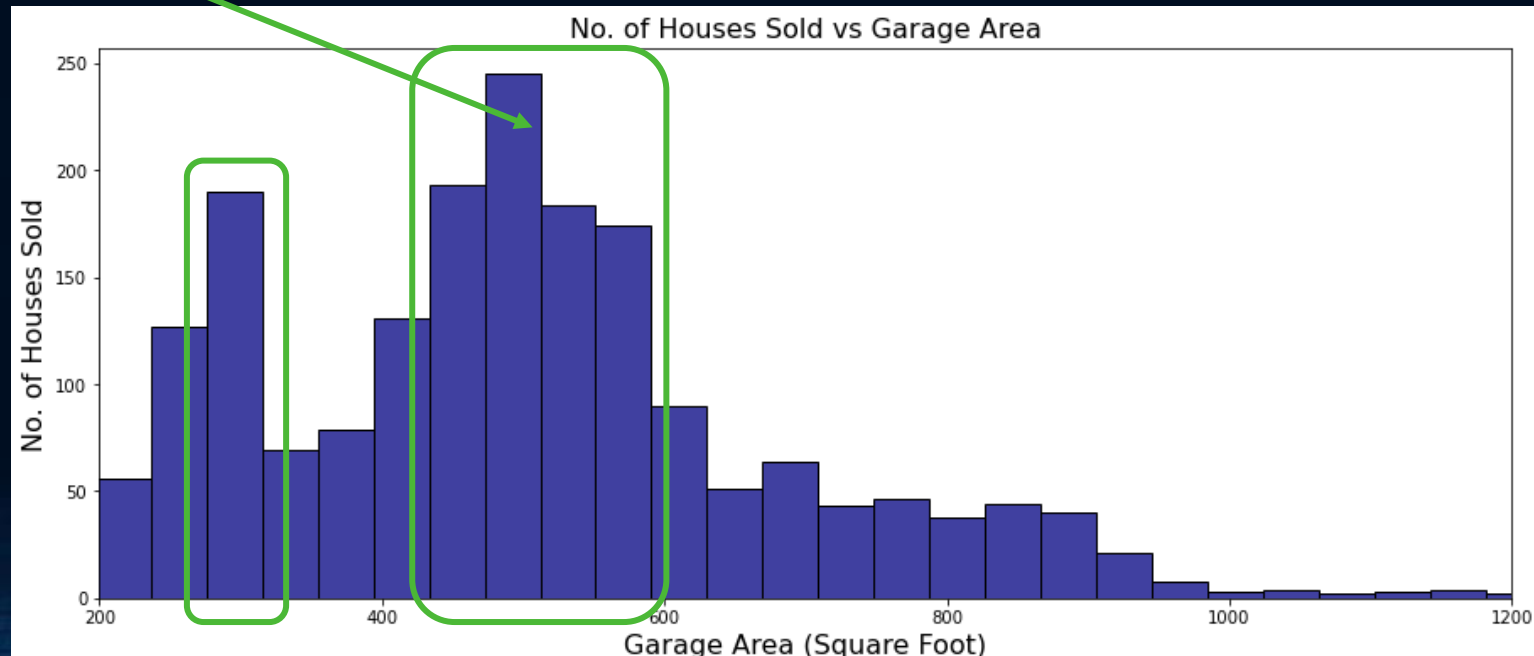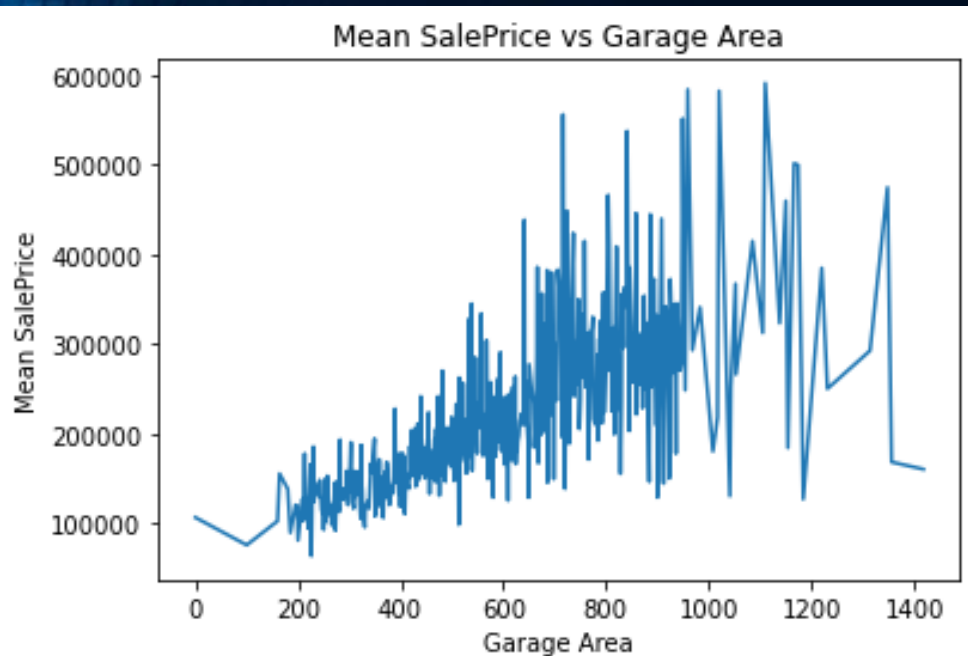
2nd Floor Square Footage

- A larger 2nd Floor Square Footage generally fetches a higher sale price.

- Imagine having a 1st Floor that's 800 - 900 sqft and...

- On top of that (literally), a 2nd Floor that measures **500 - 900 sqft**!
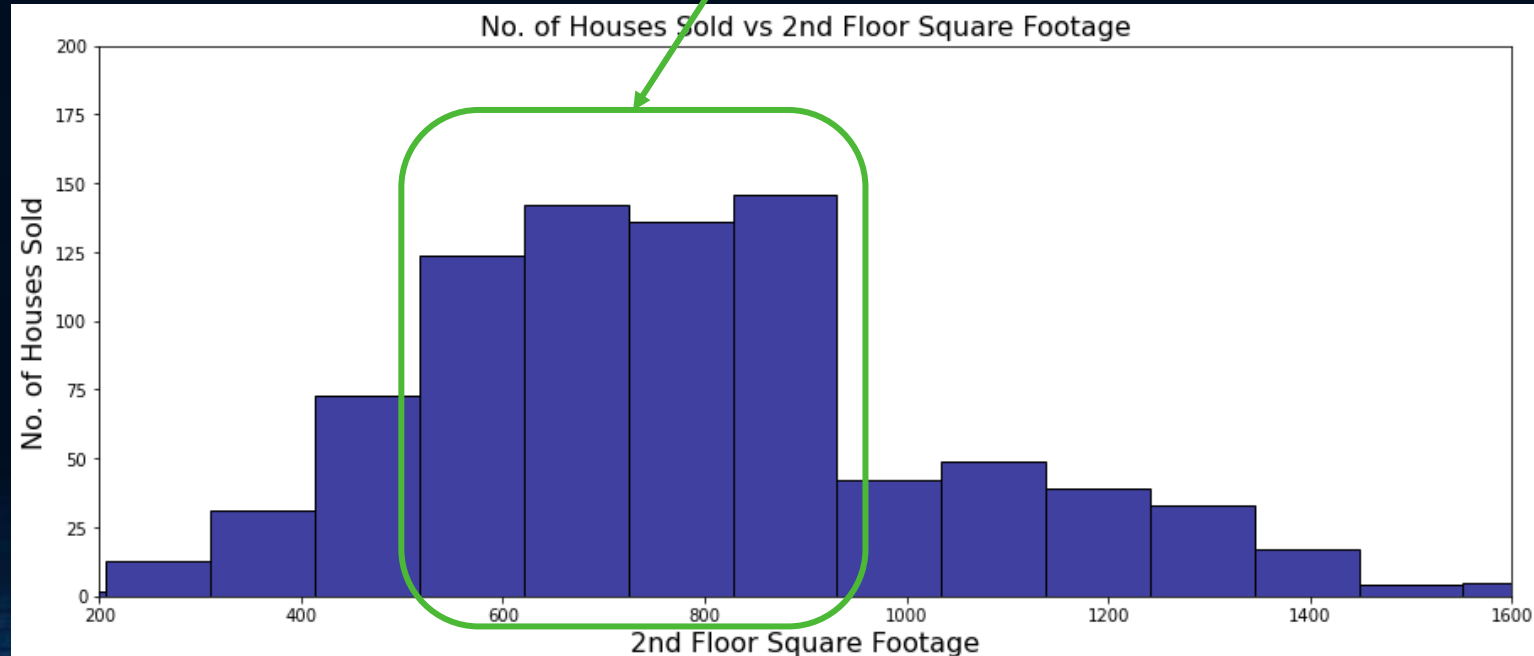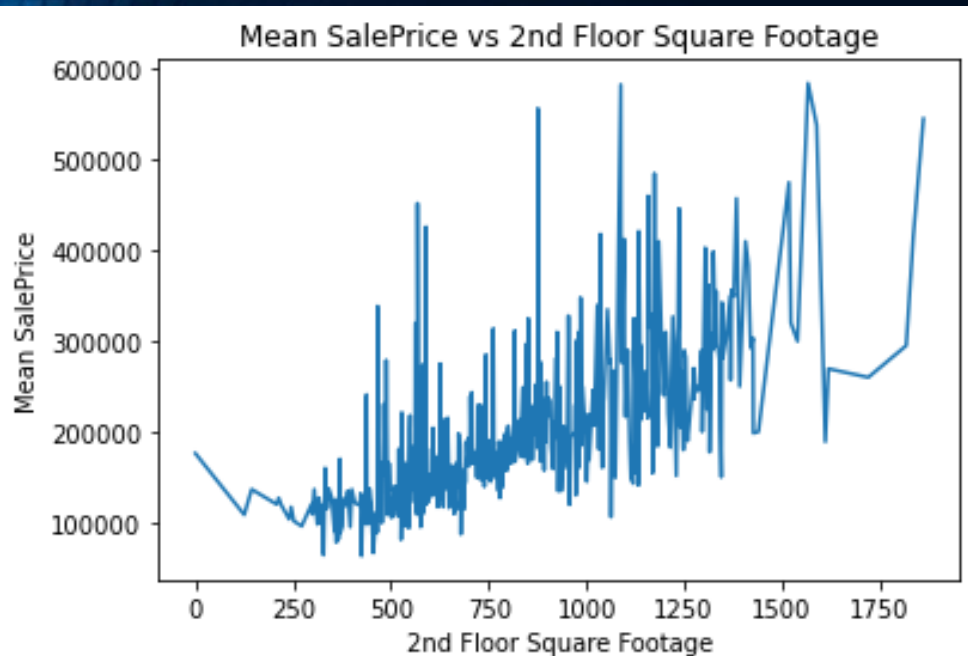
How big a Garage do you REALLY need?

- A larger Garage Area generally fetches a higher sale price.

- 300 sqft, and a range of 450 - 600 sqft seems to be the ideal garage size for most people. But why?

- The magic number is **500 sqft**.
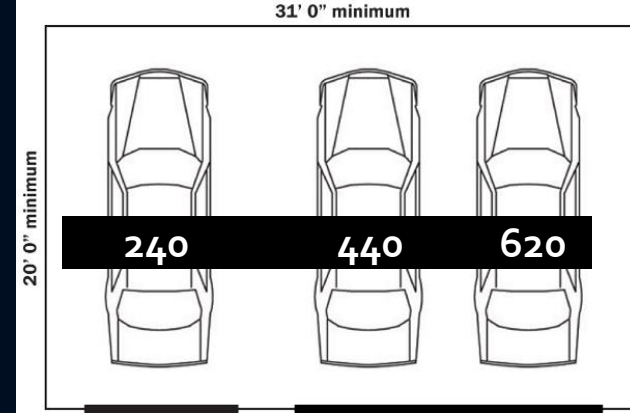
Why BIGGER is `NOT` always BETTER!
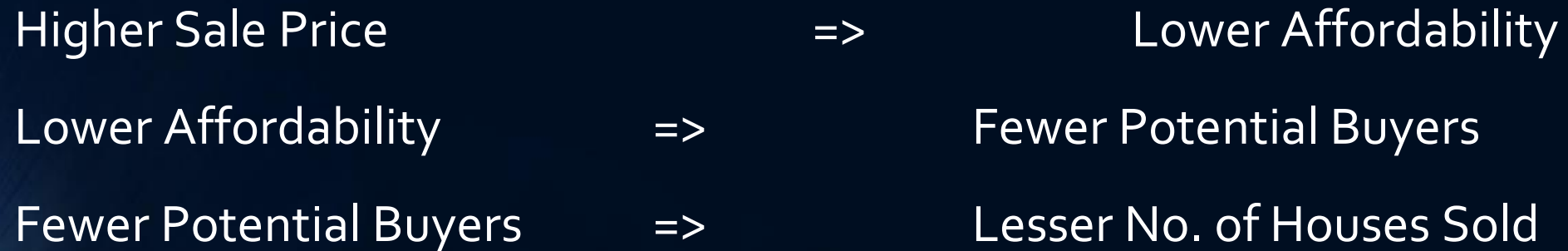
A higher square footage generally results in a higher sale price.

Higher Sale Price	=>	Lower Affordability
Lower Affordability	=>	Fewer Potential Buyers
Fewer Potential Buyers	=>	Lesser No. of Houses Sold

This is in line with our earlier data where the highest sale volume did not go to houses with the highest square footages.

Gaps that need to be filled:

Total No. of Houses in Neighborhoods
- This would allow us to evaluate the percentage % of houses that are sold across all neighborhoods.

Demographic of Neighborhood and Buyers
- Age range, Marital Status, Family size, No. of kids could play a part in helping realtors recommend the right houses to the right buyers.

Sale price in real value to take into account of inflation
- Rate of inflation would allow us to "reflect" the past sale prices based on today's economy.

Conclusion:

Neighborhood
- Focus on marketing the houses in Northridge Heights and Northridge.

Quality of House
- Advice clients with below average housing quality to refurbish their home.

Fireplaces
- Focus on marketing houses with at least 1 fireplace and encourage clients to build a fireplace if they do not have one.

Housing Exteriors
- Avoid houses with hardboard or stucco exteriors as they are prone to rot and mold especially with the above average snowfall in Ames.

BIGGER is NOT always BETTER
- A bigger house may have a higher potential sale price, but that also means that it could have a lower chance of being sold due to a lower number of potential buyers who can afford it