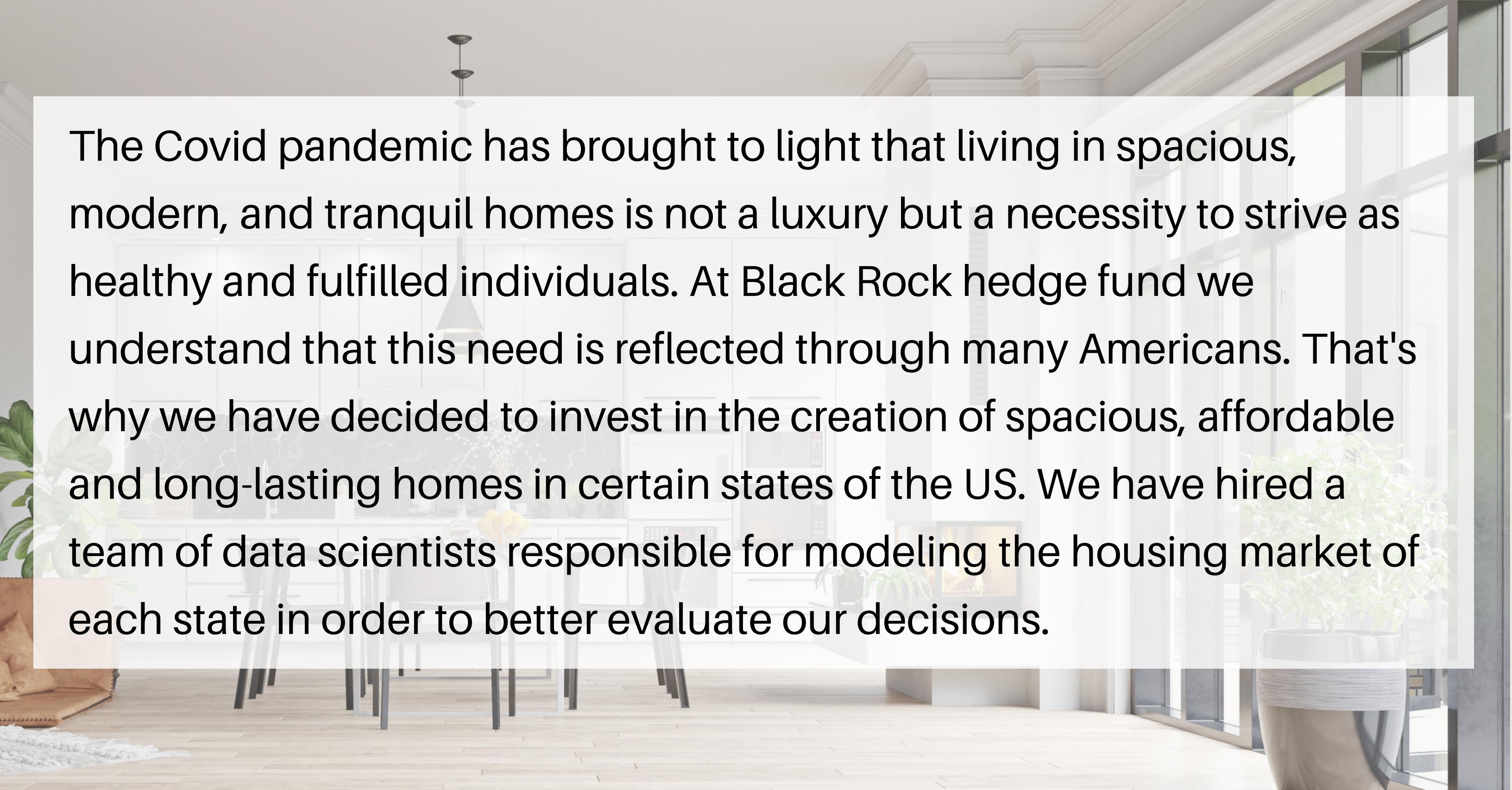




DSI 927 - Predicting sale price in Iowa's housing market

ABDELJALIL KABBAJ

MISSION STATEMENT



The Covid pandemic has brought to light that living in spacious, modern, and tranquil homes is not a luxury but a necessity to strive as healthy and fulfilled individuals. At Black Rock hedge fund we understand that this need is reflected through many Americans. That's why we have decided to invest in the creation of spacious, affordable and long-lasting homes in certain states of the US. We have hired a team of data scientists responsible for modeling the housing market of each state in order to better evaluate our decisions.

PROBLEM STATEMENT



As a data analyst responsible for modeling the housing market in Iowa, I am tasked to investigate and evaluate the highest features that will predict sale prices. Next, I investigate what people look most in homes and make recommendations on whether or not Black Rock should invest in Iowa.

Note: This is purely a thought experiment, therefore for the sake of this project, we are not taking into consideration the time frame of this model, assuming that the data listed is fairly recent.

LOOKING AT CONTINUOUS FEATURES

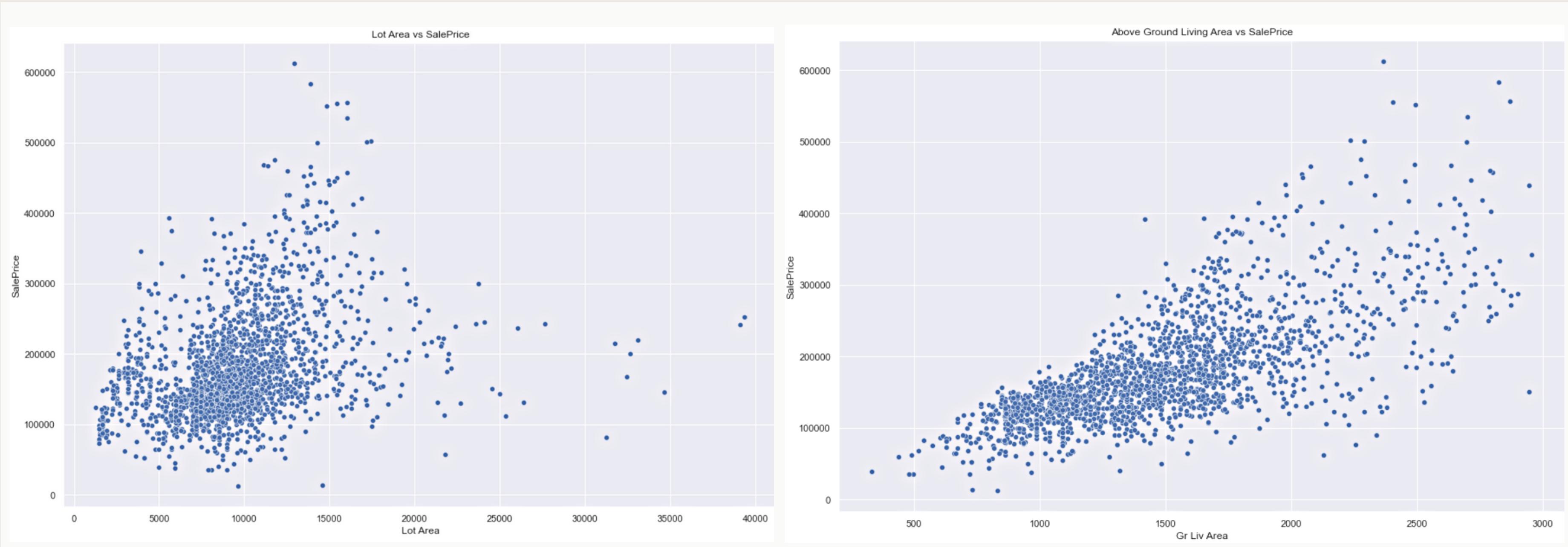


Figure 1.

Figure 2.



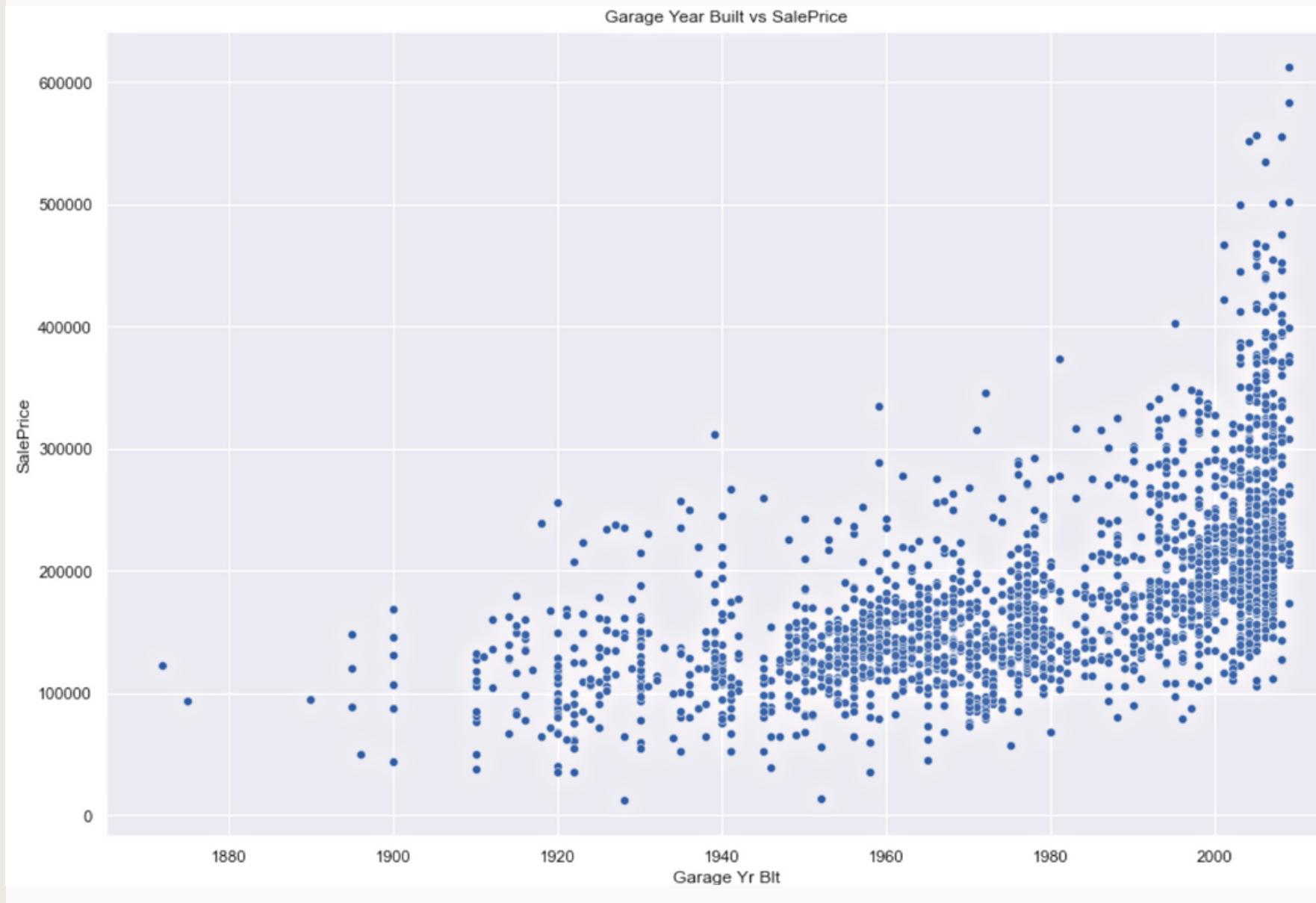


Figure 3.

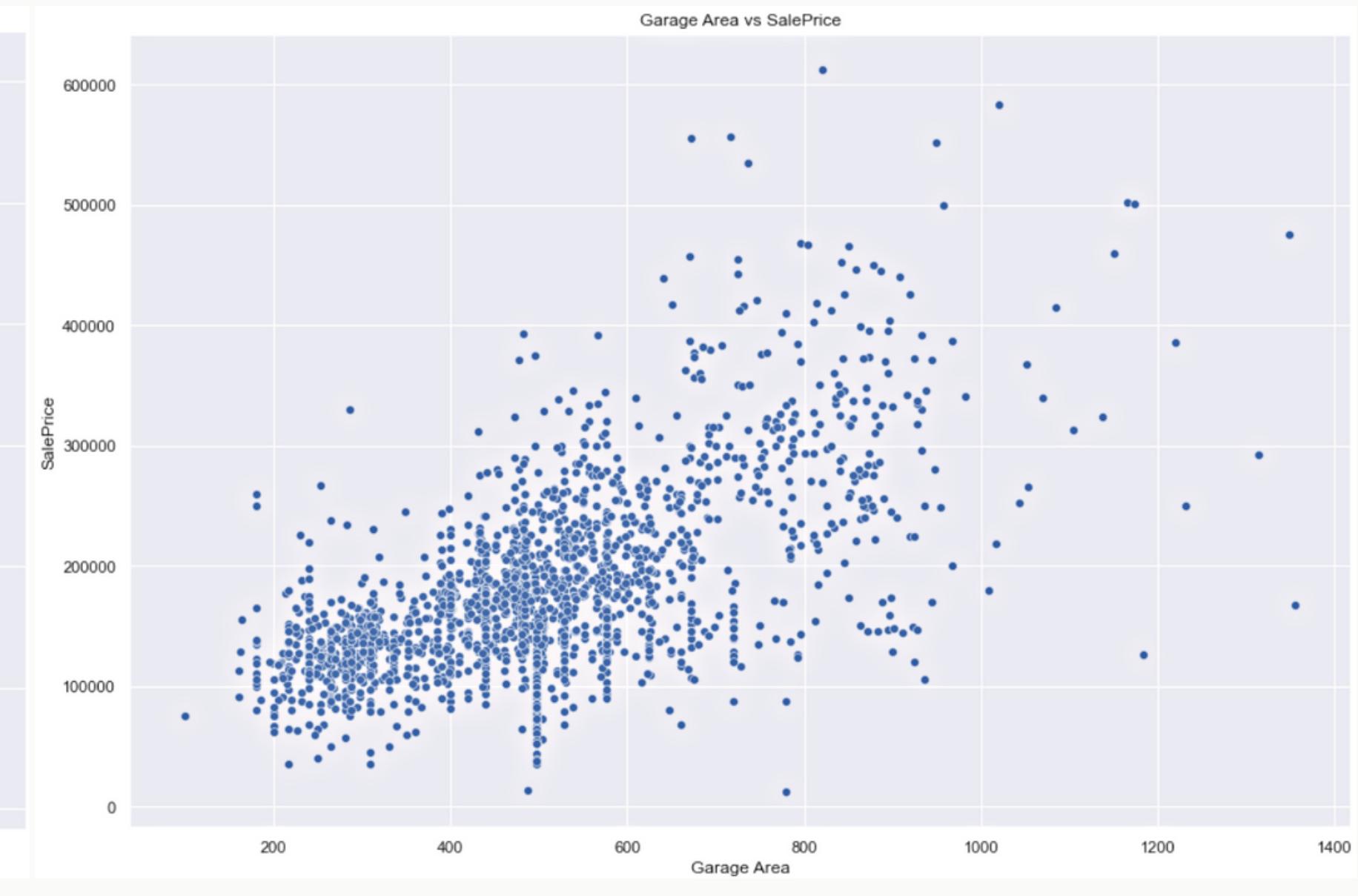


Figure 4.



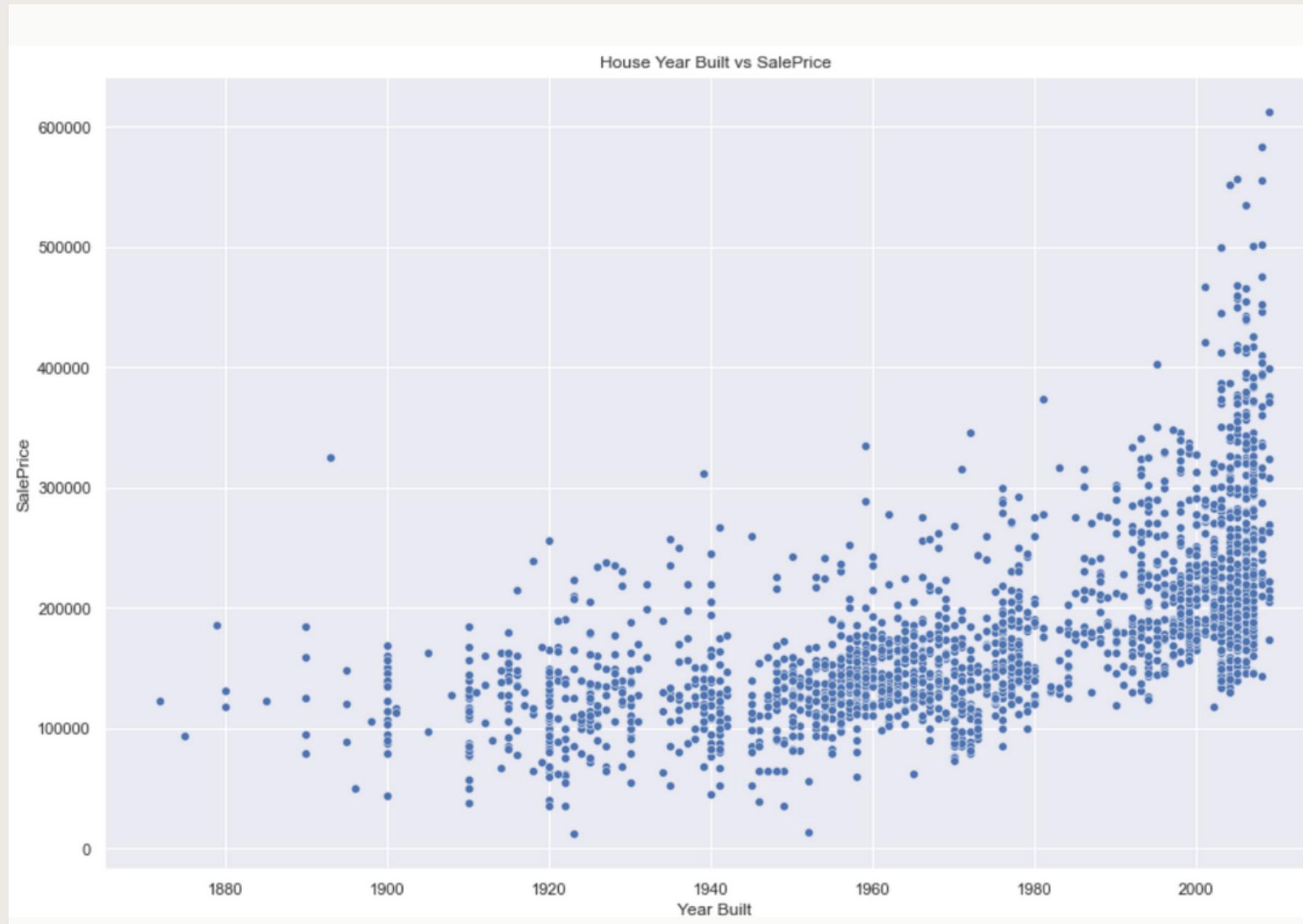


Figure 5.

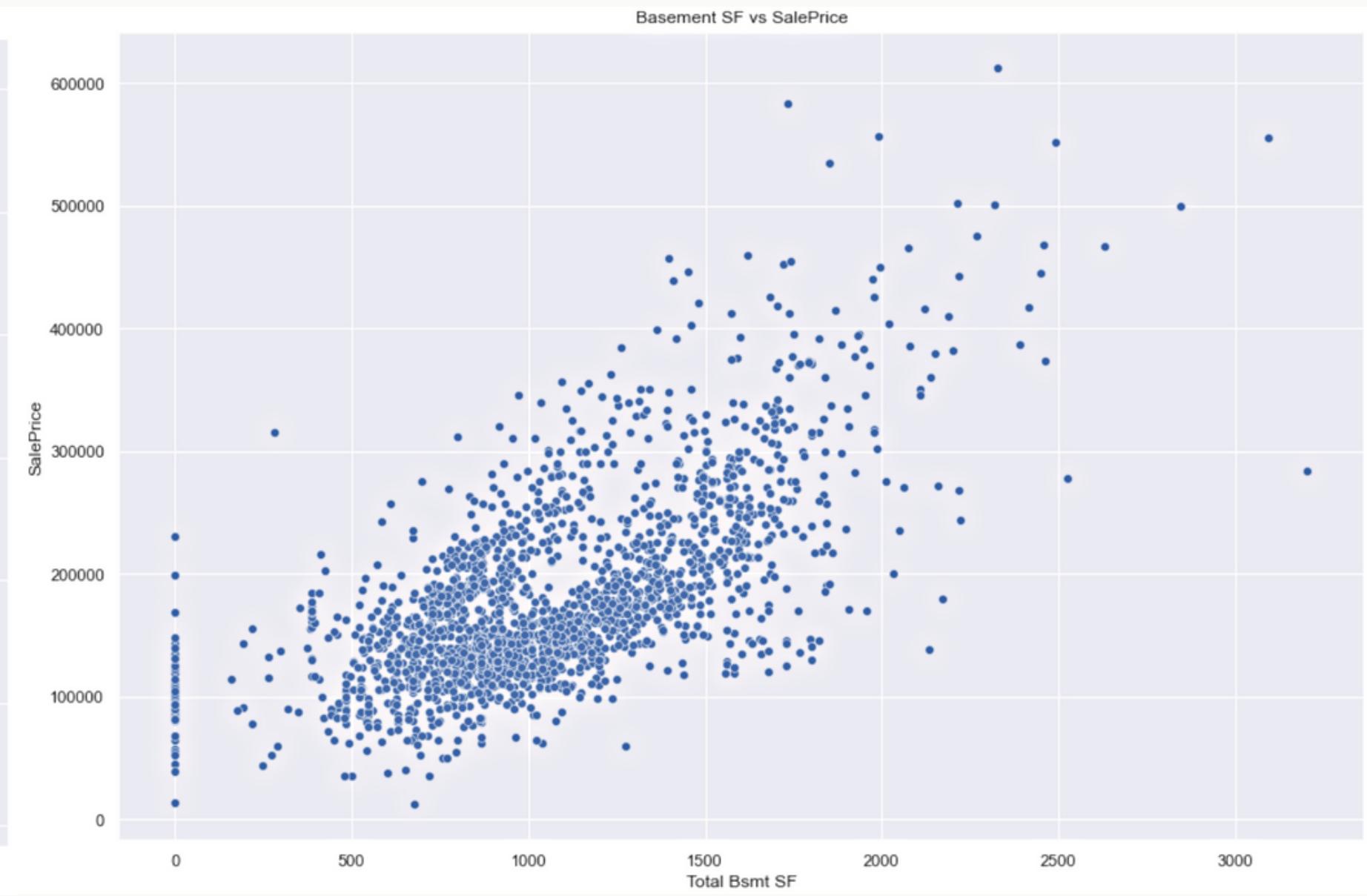


Figure 6.



LOOKING AT ORDINAL FEATURES

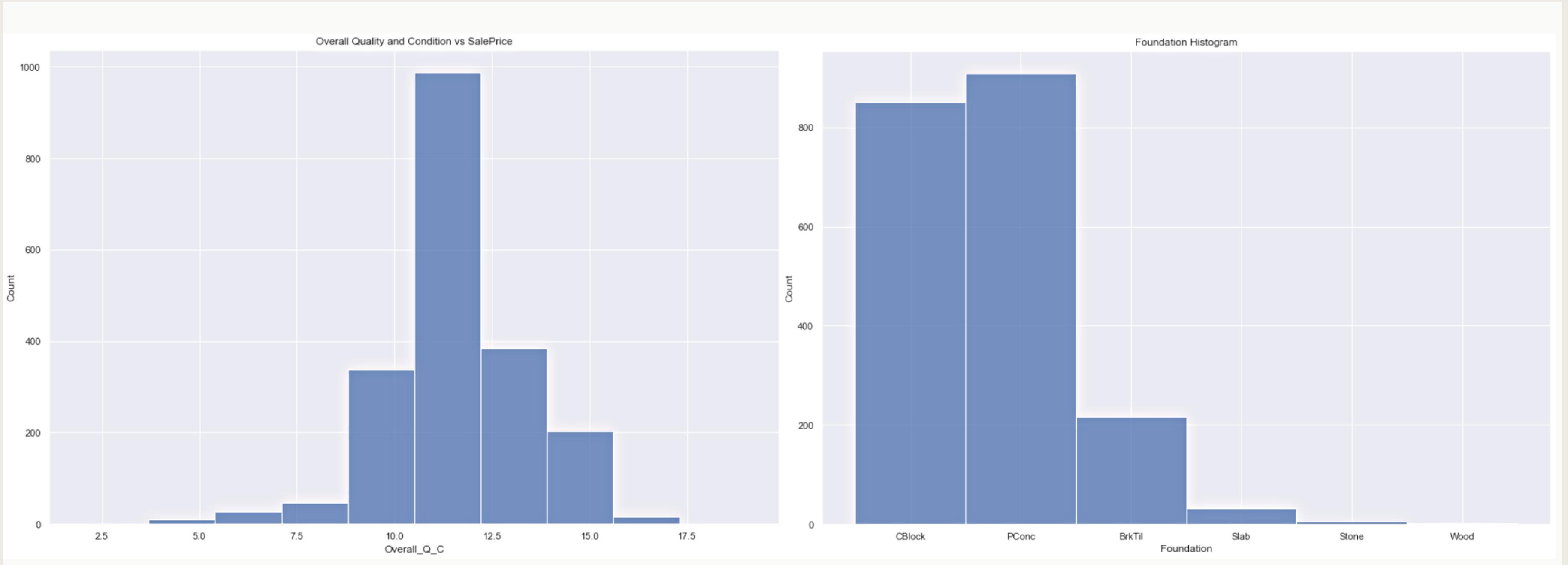


Figure 1.

Figure 2.



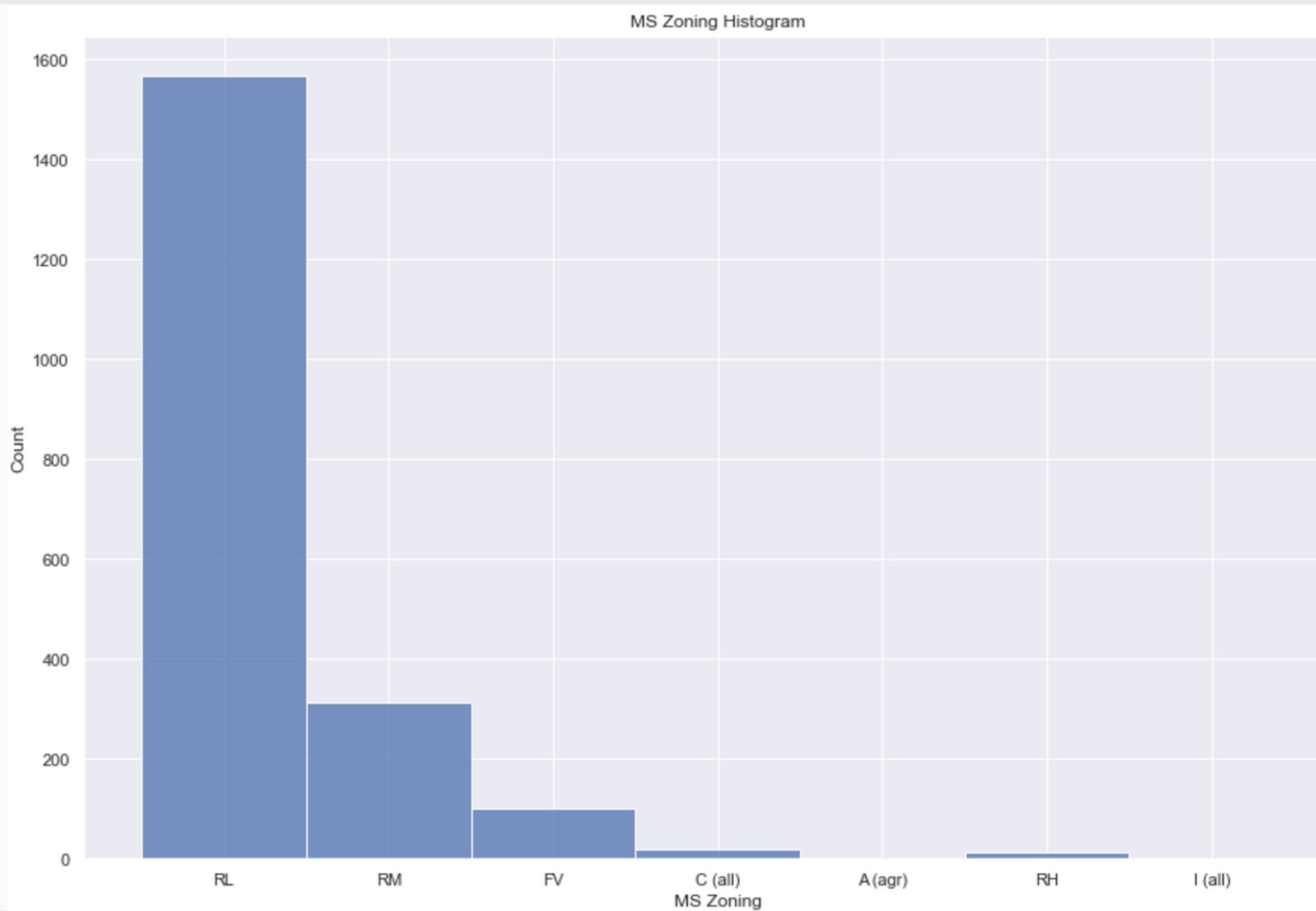


Figure 3.



Building a Linear Regression

Features I have included in my model are:

- Lot Area, Overall Quality and Condition, Year Built, Above Grade Living Area, Garage Area, Total Rooms Above Grade, First and Second Floor Combined Square Footage, Neighborhood, Total Baths, Foundation, MS Zoning, Garage Type, Basement Square Footage

Features I have dummified:

- Total Rooms Above Grade, Year Built, Total Baths, Foundation, MS Zoning, Garage Type, Neighborhood



Results of Linear Regression

Cross validation score :

0.8636

Train Score:

0.8944

RMSE after split:

28753.1164

RMSE before split:

25021.9579

Test Score:

0.8664



EVALUATING MY LINEAR REGRESSION

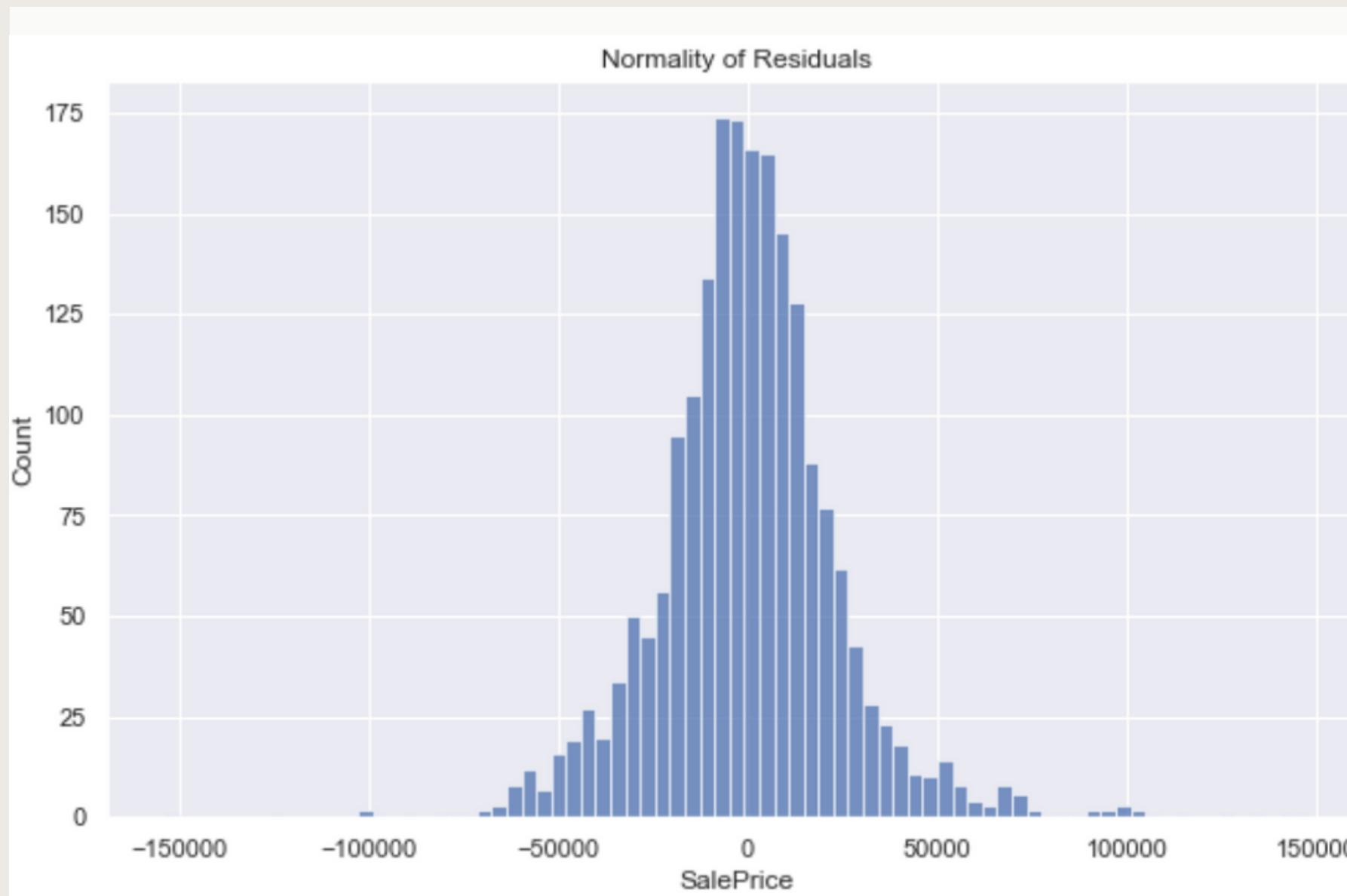


Figure 1.

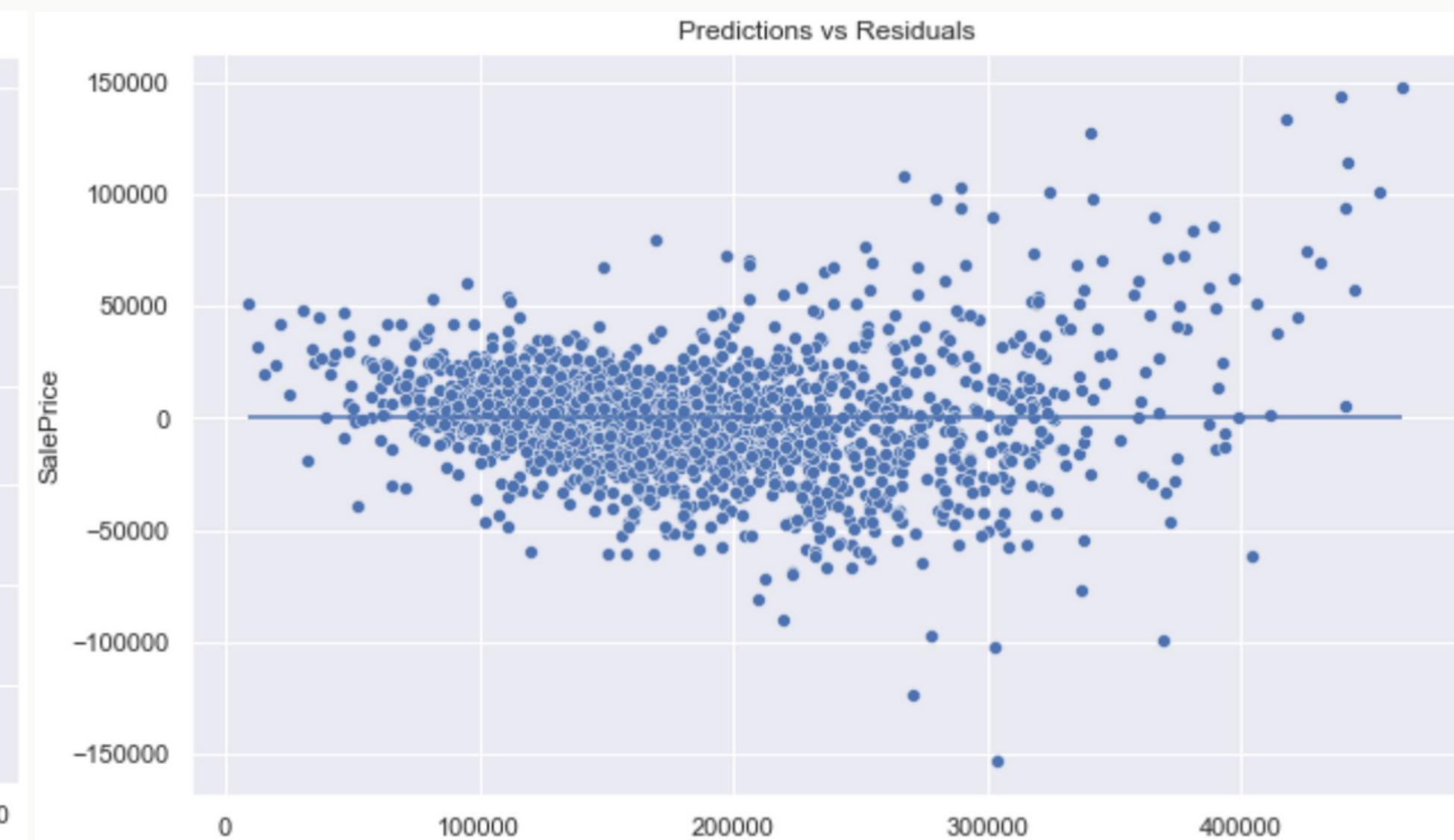


Figure 2.



Results of regularizing my model through Ridge

Train Score:

0.89202

RMSE:

28571.4014

Test Score:

0.8681



Results of regularizing my model through Lasso

Train Score:

0.8944

RMSE:

28681.1024

Test Score:

0.8671



INTERPRETING COEFFICIENTS - CONTINUOUS

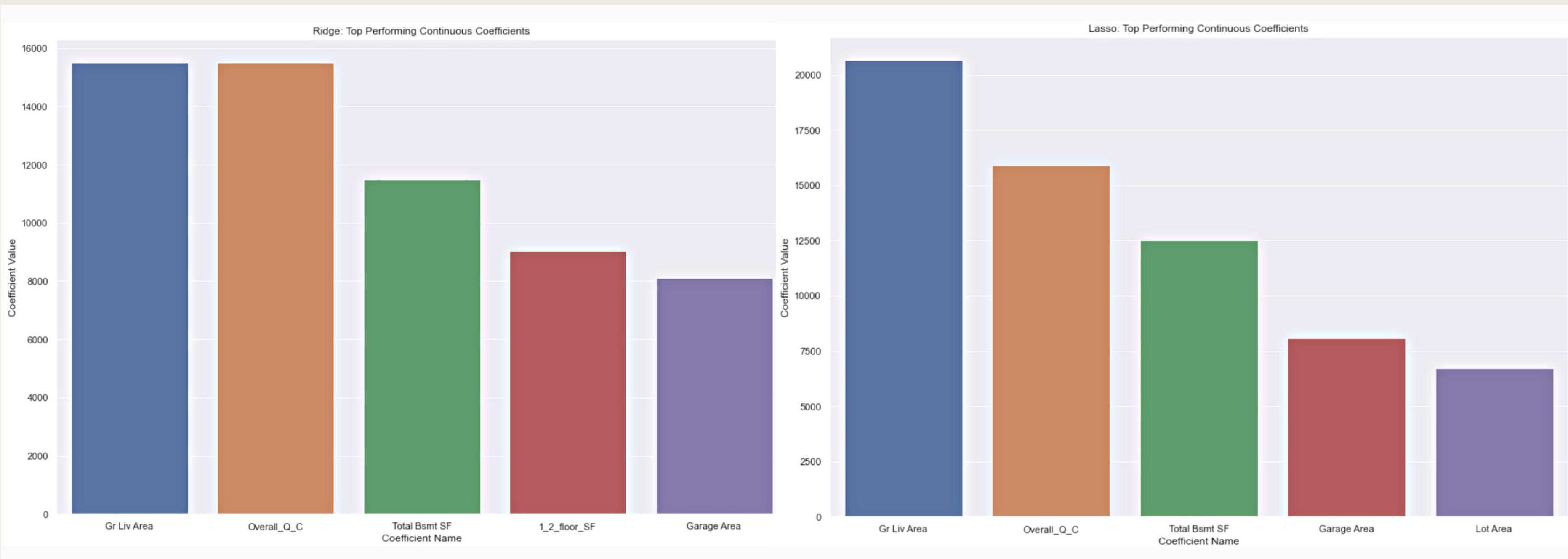
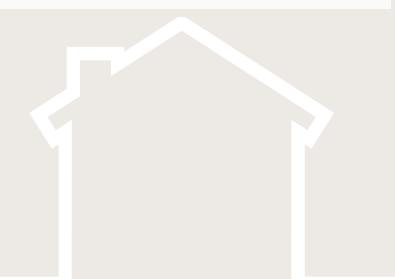


Figure 1.

Figure 2.



INTERPRETING COEFFICIENTS - MS ZONING

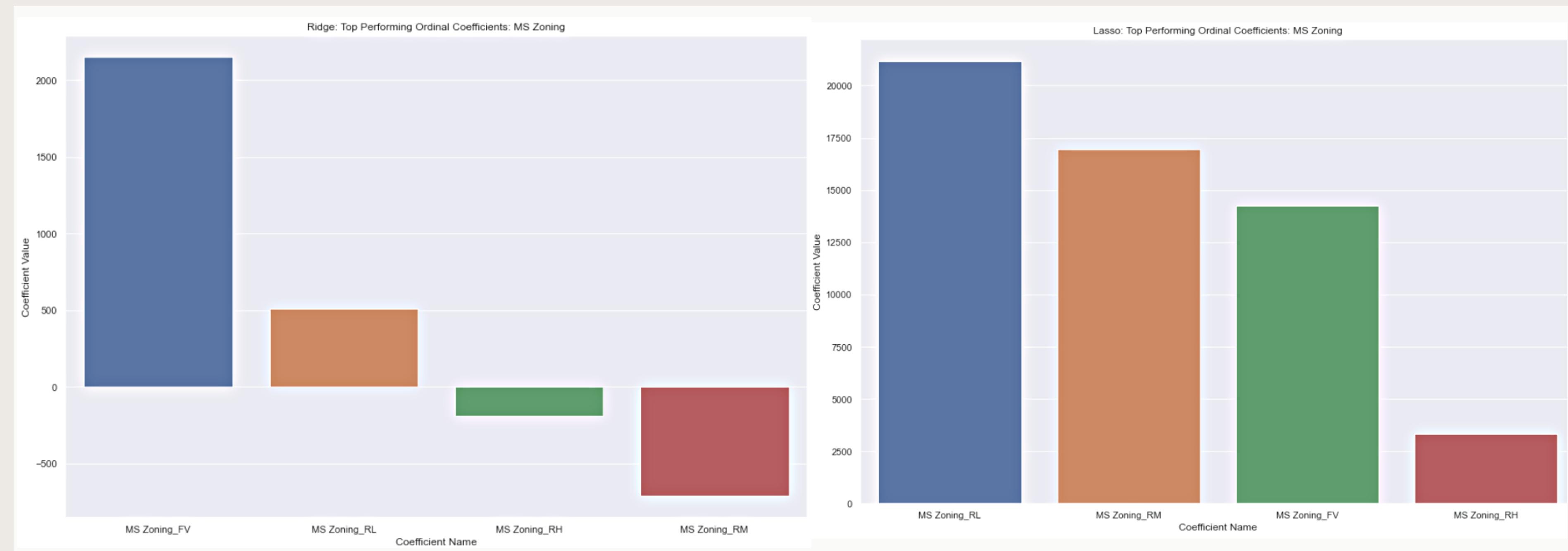


Figure 1.

Figure 3.



INTERPRETING COEFFICIENTS- YEAR BUILT

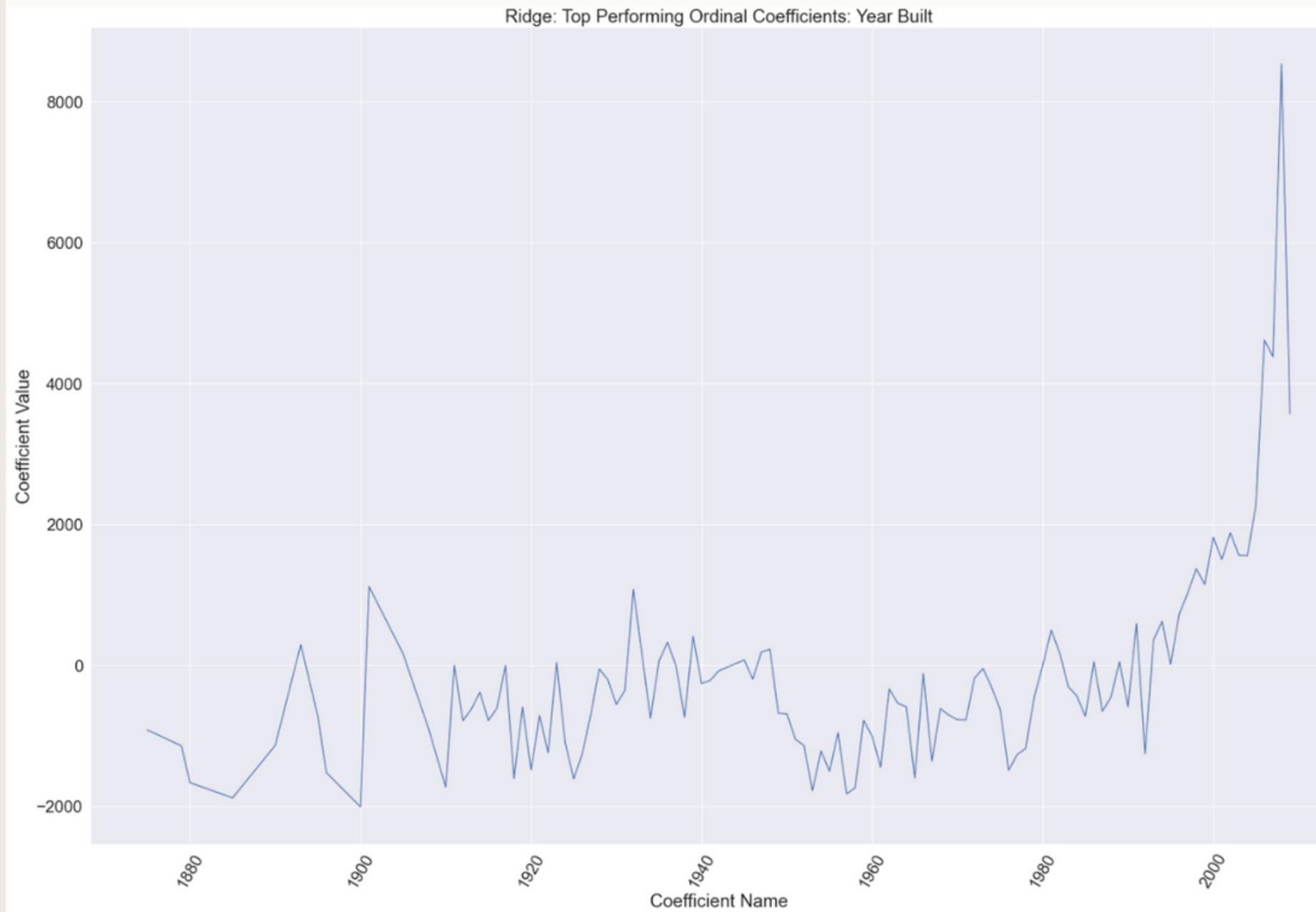


Figure 3.

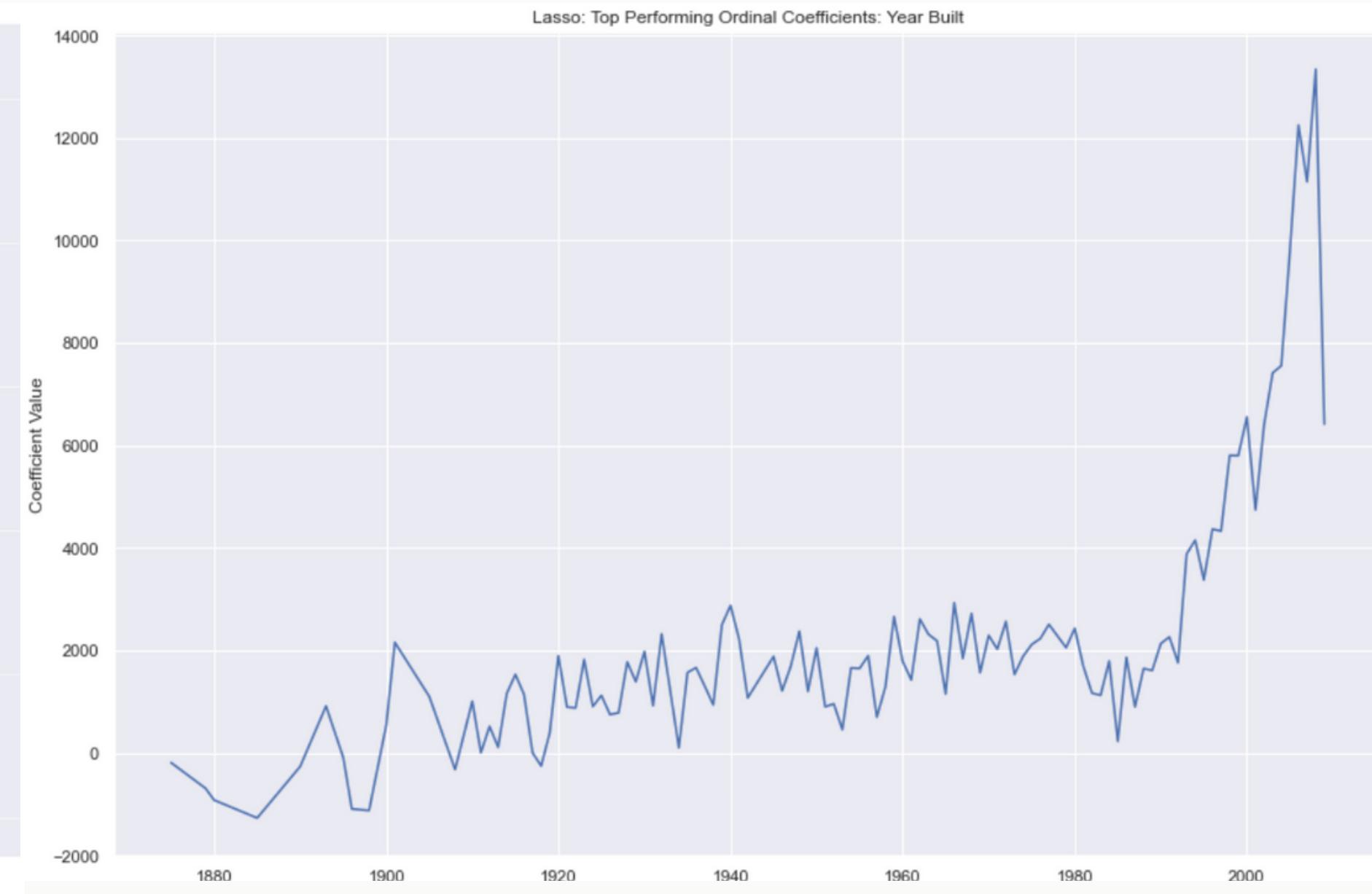


Figure 4.



INTERPRETING COEFFICIENTS- GARAGE TYPE

Ridge: Top Performing Ordinal Coefficients: Garage Type

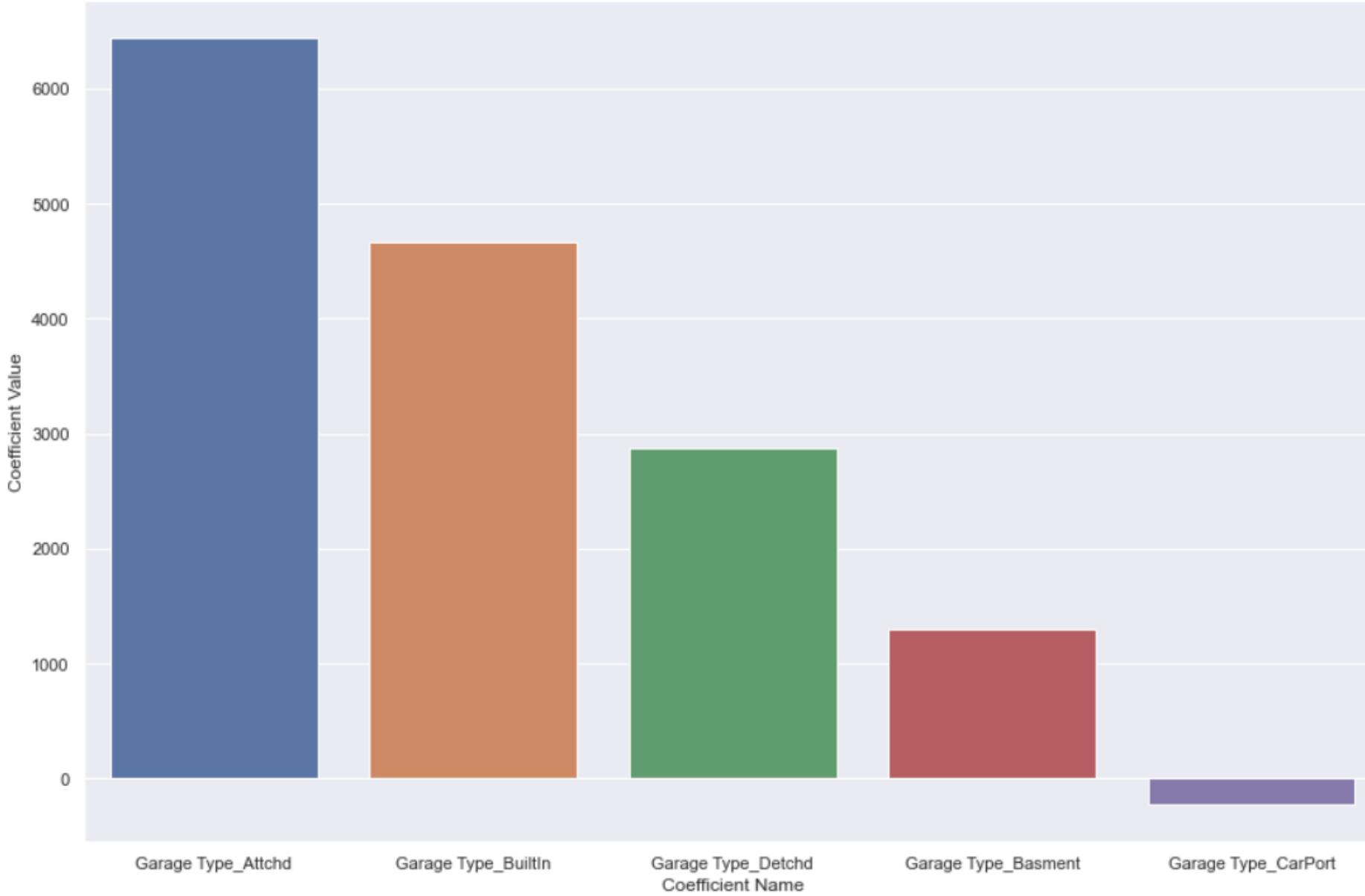


Figure 1.

Lasso: Top Performing Ordinal Coefficients: Garage Type

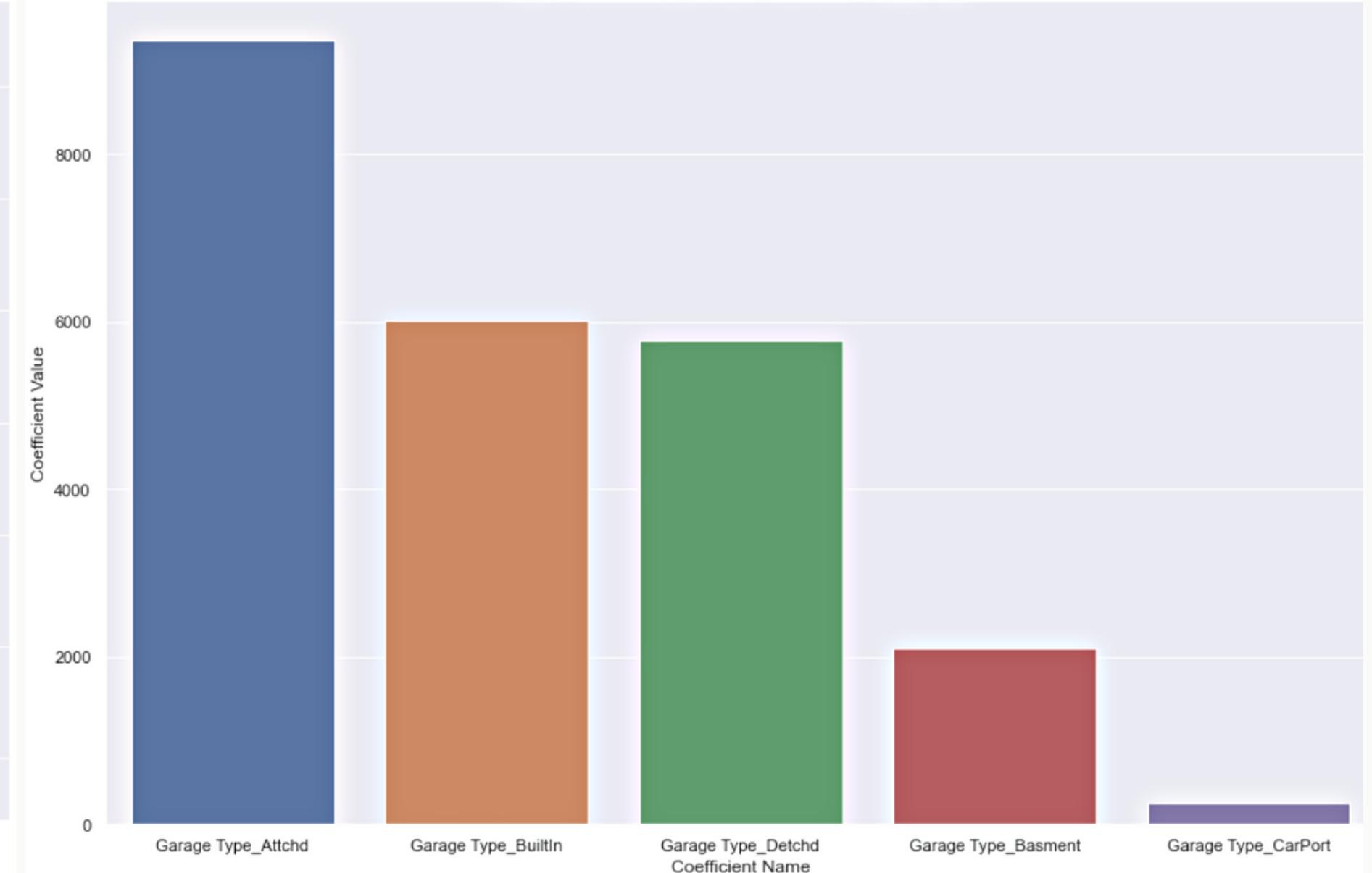


Figure 2.



Summary of Results and Recommendations

- I believe that the state of Iowa is qualified for our investment, as we are looking to build spacious, tranquil, and modern homes.
 - We can infer the following:
 - People in Iowa enjoy indoor spaces.
 - People in Iowa want accessibility when looking for a garage.
 - People in Iowa enjoy living in less crowded densities.
 - People in Iowa look for recently built houses.



Thank
You.
Do you have any
questions?

ABDELJALIL KABBAJ

