

KINGS COUNTY REAL ESTATE INSIGHTS

RITTENHOUSE BROTHERS



INTRODUCTION

This study aims to address the importance of identifying the factors that significantly influence home costs. Because conventional methods rely on anecdotal evidence or limited research, they often yield incorrect findings. This knowledge gap makes it more difficult for stakeholders to predict and assess changes in house prices. To solve this issue, we are using multiple regression modeling techniques to thoroughly analyze home sales data. Multiple regression is a useful tool for analyzing the relationship between different attributes and home sales prices since it takes into consideration the combined impact of several independent variables.



PROBLEM STATEMENT

Real estate valuation poses significant challenges for Rittenhouse Brothers, particularly when determining property values influenced by unique features and the impact of renovations or upgrades. The subjective nature of valuation, varying interpretations among appraisers and professionals, and the absence of a purely objective methodology contribute to valuation discrepancies.



MAIN OBJECTIVE

This Analysis aims to establish an objective property valuation model, focusing specifically on the Property's Unique features and the impact of renovations or upgrades. The primary goal is to minimize variations in valuation estimates and provide Rittenhouse Brothers with a more standardized and reliable method for assessing property values



SPECIFIC OBJECTIVES

Perform exploratory data analysis to uncover connections between various variables and the target variable. This process aids in identifying pertinent variables for inclusion in a regression model.

Create a multiple regression model to forecast house sale prices by taking into account chosen independent variables and examining their influence on the dependent variable. Validate the model assumptions, evaluate its fitness for the data, and refine the model as needed.

Analyze the coefficients of the independent variables within the model to discern their individual contributions to house prices. Identify the most impactful factors influencing house sale prices and delineate their respective effects.

Assess and confirm the efficacy of the model.

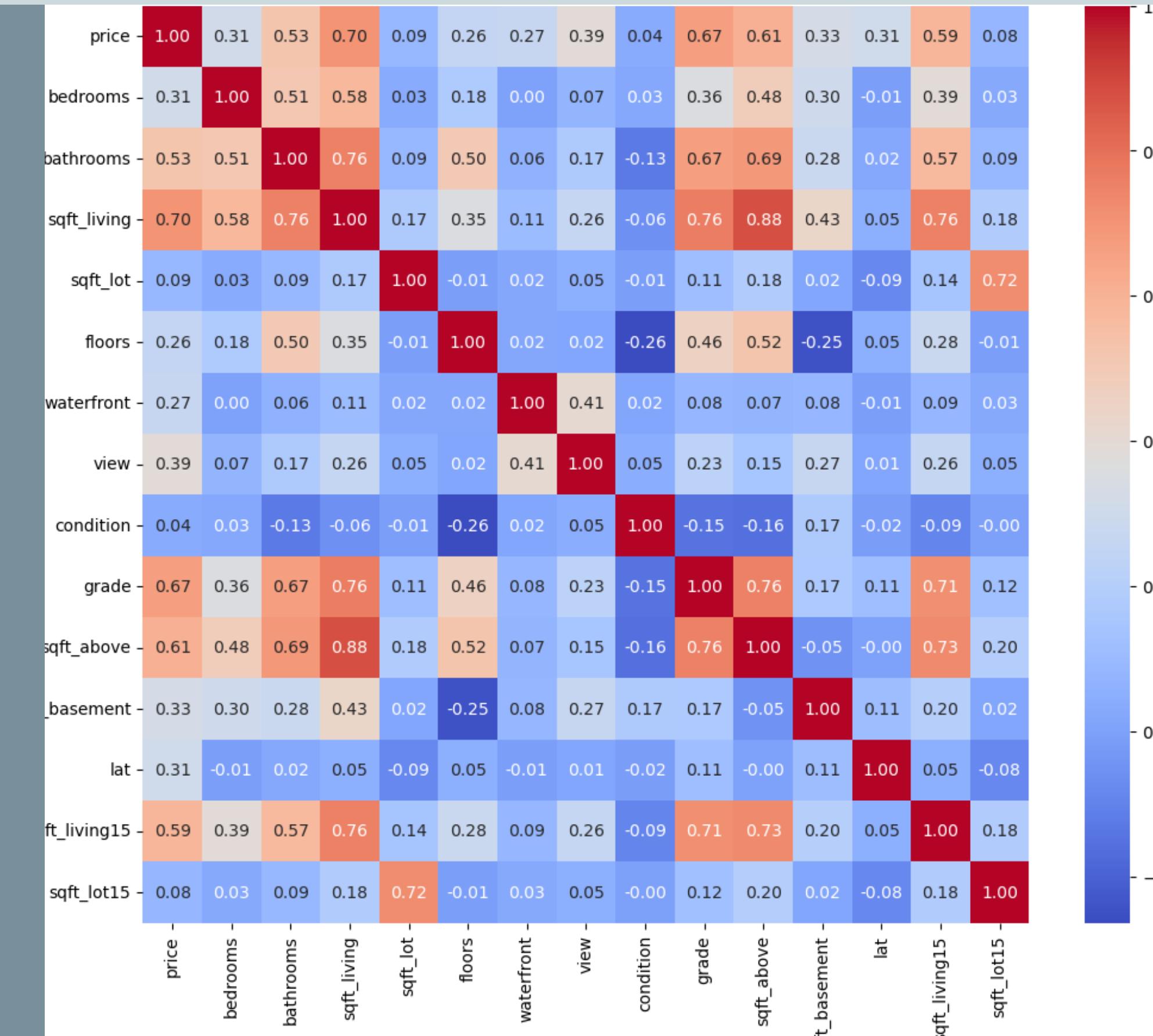
Offer practical insights and recommendations derived from the analysis to aid Rittenhouse Brothers in making well-informed decisions related to property investment

EXPLORATORY DATA ANALYSIS

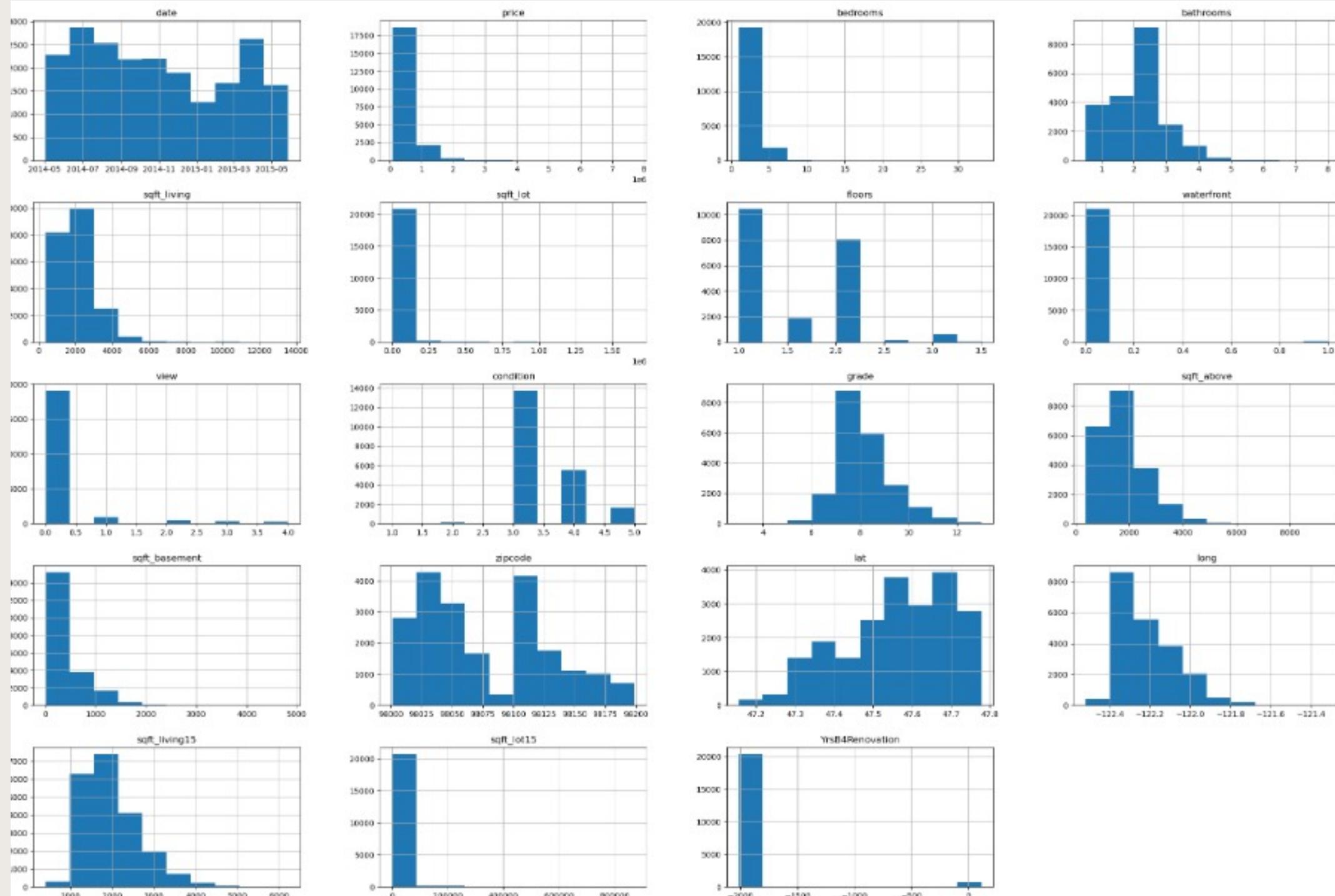
CORRELATION MATRIX HEATMAP

From this, we can see there is a connection between square feet living and the number of bathrooms and the number of bedrooms

The square footage of a house's living space closely resembles that of its 15 nearest neighbors, indicating a sense of neighborhood similarity.



DISTRIBUTION OF PREDICTOR VARIABLES



This shows the distribution of predictor variables

REGRESSION MODEL

For this project, we created a multivariate linear regression model that uses a variety of independent variables to forecast house selling prices. The following were the primary independent variables in the model: "sqft_living," "floors," "lat," "zipcode," and "views." We handled missing values, outliers, and feature scaling by performing data cleaning and preprocessing procedures before model fitting.

```
print(model.summary())
```

OLS Regression Results						
Dep. Variable:	price	R-squared:	0.672			
Model:	OLS	Adj. R-squared:	0.672			
Method:	Least Squares	F-statistic:	8673.			
Date:	Tue, 02 Jan 2024	Prob (F-statistic):	0.00			
Time:	19:53:22	Log-Likelihood:	-4649.1			
No. Observations:	21143	AIC:	9310.			
Df Residuals:	21137	BIC:	9358.			
Df Model:	5					
Covariance Type:	nonrobust					

	coef	std err	t	P> t	[0.025	0.975]

const	13.0484	0.002	6292.472	0.000	13.044	13.052
sqft_living	0.7181	0.006	128.021	0.000	0.707	0.729
floors	0.0692	0.004	16.656	0.000	0.061	0.077
lat	1.6340	0.016	104.540	0.000	1.603	1.665
zipcode	-0.0004	4.19e-05	-8.910	0.000	-0.000	-0.000
view	0.1508	0.003	48.050	0.000	0.145	0.157

Omnibus:	387.036	Durbin-Watson:	1.994			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	504.870			
Skew:	0.246	Prob(JB):	2.34e-110			
Kurtosis:	3.576	Cond. No.	404.			

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified

Since only price, i.e. the target variable, and sqft_living has been log-transformed, the model indicates that a one-unit increase in floors leads to a 9.1% increase in house prices. A one-level improvement in condition (Cond. No.) leads to a 404 times increase in the standard error of the estimate.

RESULTS

- "For every 1% increase in the size of the living area (square footage), we expect to see, on average, a 0.84% increase in the price of the house. So, a larger living space tends to be associated with a proportionate increase in house prices."
- "Considering the number of floors, each additional floor in a house is associated with a substantial increase. Specifically, a one-floor increase is linked to an estimated 9.4% rise in house prices. This highlights the impact that additional levels can have on the overall value of a property."
- "Improving the condition of the house by one level is significant. It's estimated that a one-level improvement in house conditions is correlated with a 7.34% increase in house prices. This underscores the value of enhancing and maintaining a property's overall condition."
- "The view of a property also plays a crucial role. A one-level improvement in the view is associated with an estimated 14% increase in house prices. A better view tends to contribute significantly to a home's overall appeal and value."



RECOMMENDATIONS

1. Importance of Geographical Location:
 - Our initial analysis has identified that the geographical location of a house plays a significant role in determining its price. However, to provide a more comprehensive understanding, we need to delve deeper into how specific locations impact housing prices. This will involve exploring factors such as neighborhood characteristics, proximity to amenities, and local market trends
2. We are continuously refining and evolving our model to ensure our model performs optimally across the entire spectrum of house prices



LIMITATIONS

1. When we're dealing with more expensive homes, the accuracy of our appraiser could be more precise.
2. Our model doesn't account for external factors like market trends, changes in regulations, or broader economic influences. The model is focused on the intrinsic qualities of properties and their enhancements.
3. Latitude and zipcode have emerged as influential factors affecting house prices. Yet, to fully understand this relationship, we would need additional context about the geographic regions in our dataset. This includes details about the specific areas or neighborhoods, which are crucial for interpreting the impact of these predictors



THANK YOU!