

PHASE 2 PROJECT



King County House Sales Analysis

BY:

Angella Bor | Jackline Njuguna
Vitelis Siocha | Mwenda Mugambi





Project Overview

Weichert Realtors needs to provide valuable advice to homeowners regarding how home renovations may impact the estimated value of their homes and by what amount.

Therefore, this analysis aims at building a model of determinants of house prices in the northwestern county and provide actionable insights for **Weichert Realtors** that will help when advising homeowners.

Business Problem

In the competitive northwestern county real estate market, **Weichert Realtors** and **Homeowners** lack clarity on which property features most significantly influence house prices. There's a need to identify these critical determinants to guide strategic renovations, pricing decisions, and investment strategies, ensuring maximum ROI and market competitiveness.



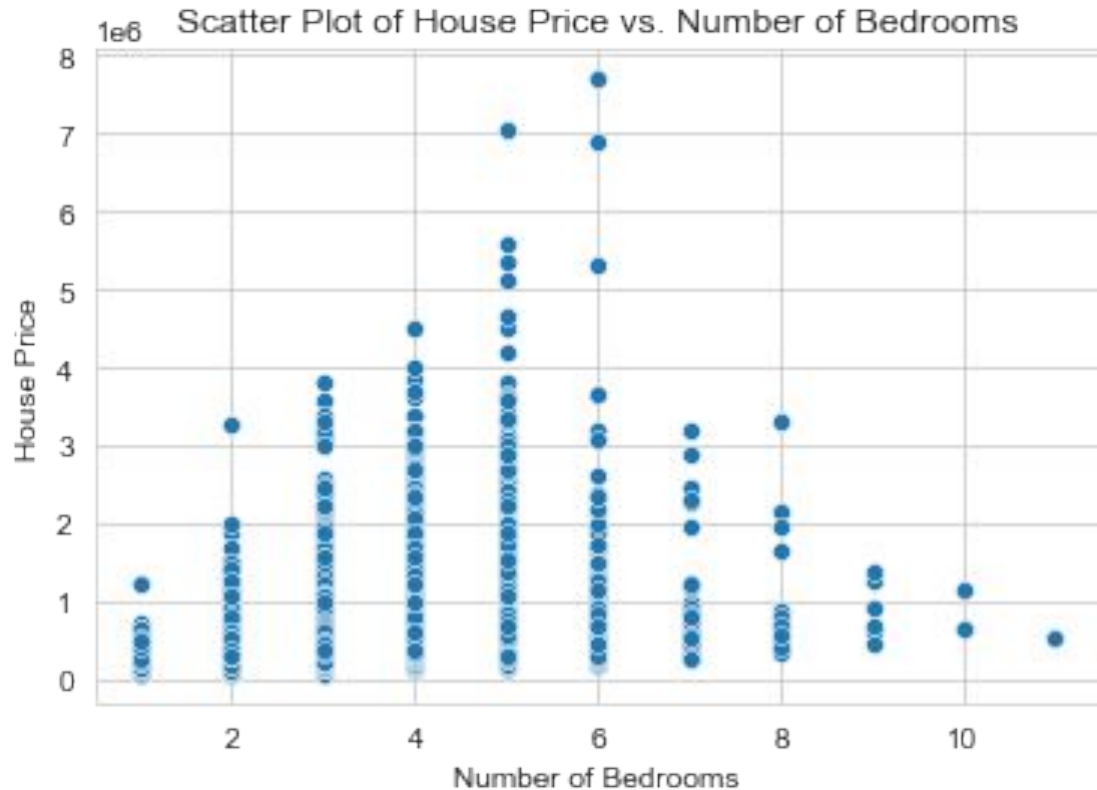
Data Understanding and EDA

We used **Exploratory Data Analysis** and **Modeling** on the **King County House Sales dataset**, which had **21597** rows and **21** columns.

Out of the 21 columns, we used the following columns: price, bedrooms, bathrooms, sqft_living, sqft_lot, floors, waterfront, view, condition, grade, sqft_above, sqft_basement. The rest of the columns were dropped.

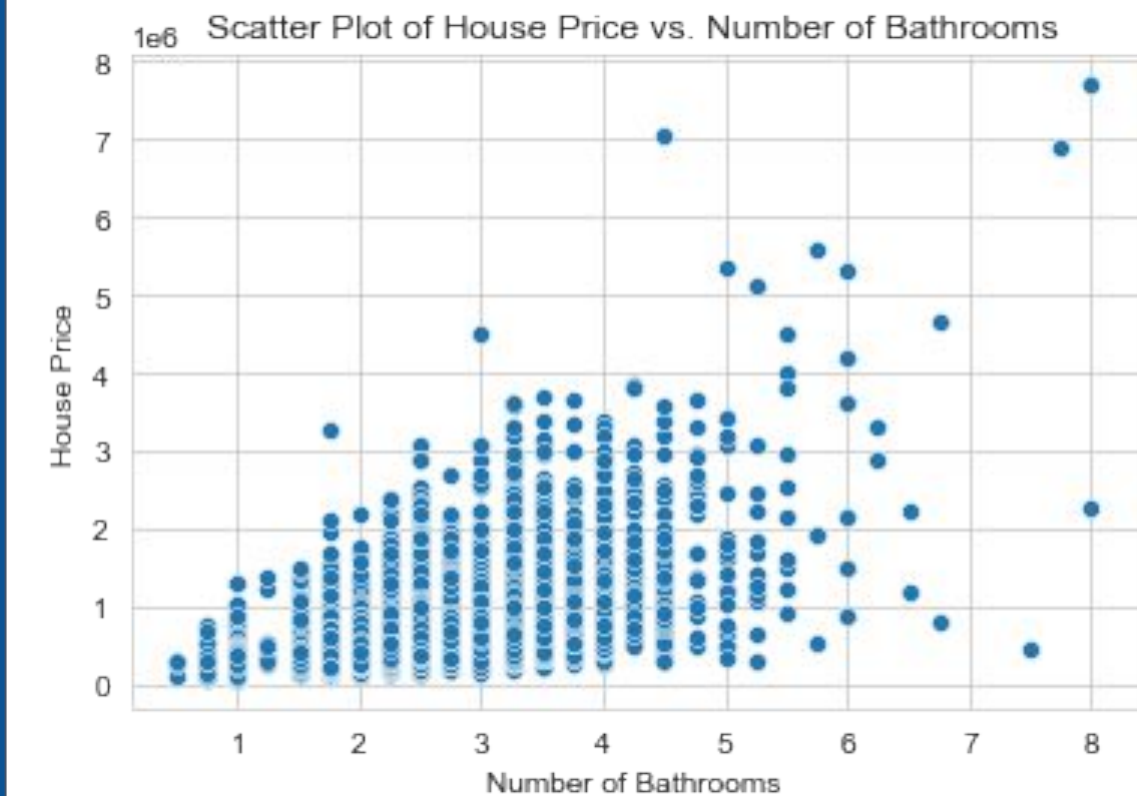


Relationship Between Price and Number of Bedrooms



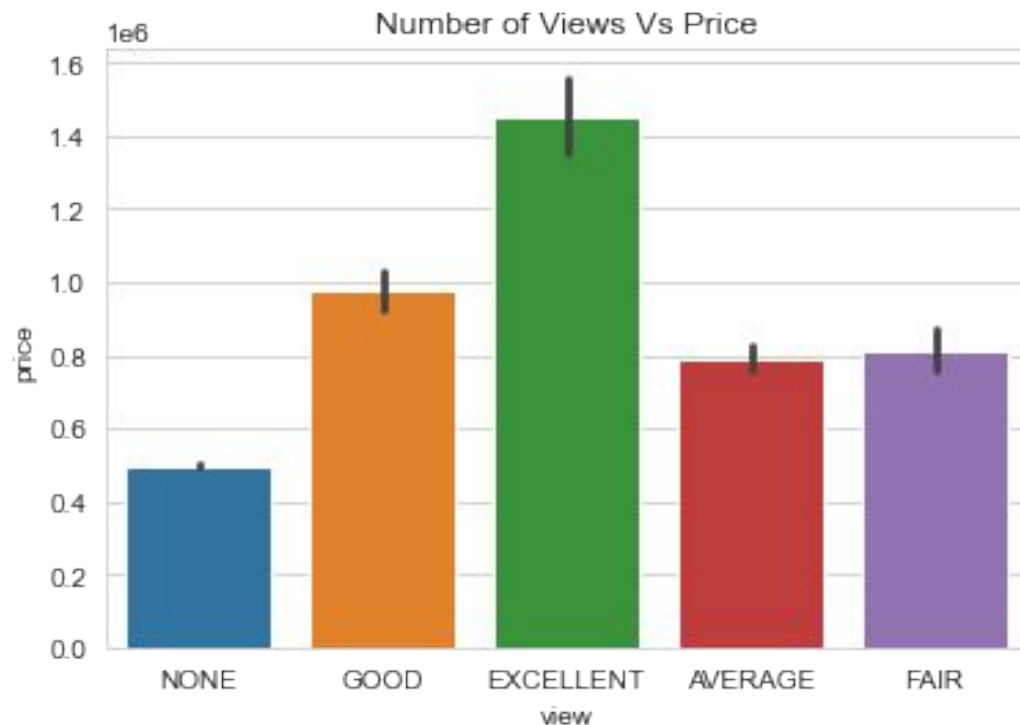
The price of the house increases with the increase in the number of bedrooms. However as the number of bedrooms reach 6, the price starts to fall.

Relationship Between Price and Number of Bathrooms



More bathrooms in a house are linked to higher selling prices."

Relationship between house views and price



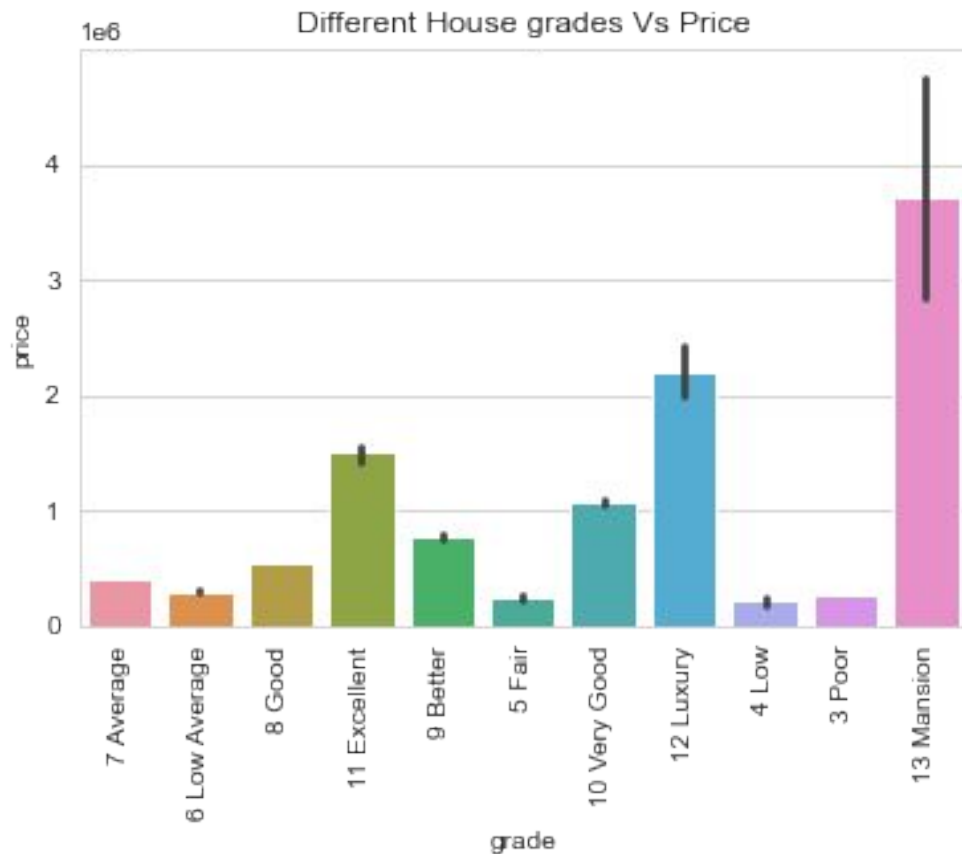
Houses with good rating views tend to have higher or better prices.

Relationship Between House condition and price



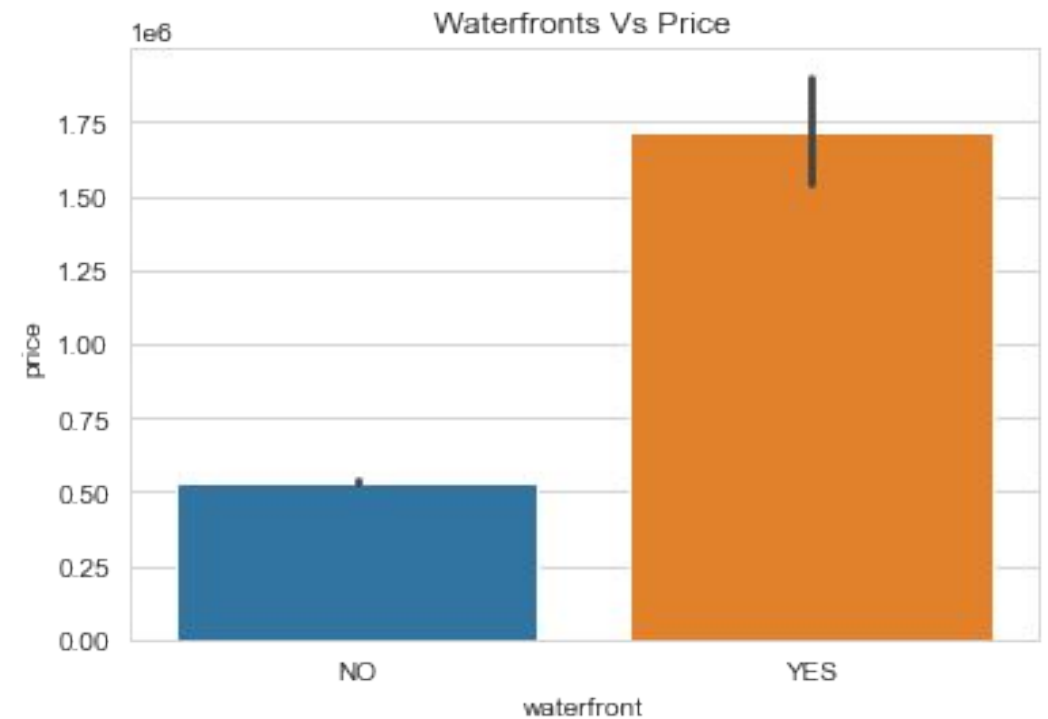
Houses in better conditions are sold at higher prices

Relationship Between grade of the house and price



Houses graded as "Mansion," and "Luxury," exhibit the highest prices.

Relationship Between House condition and price



Houses with waterfronts have extremely high prices

Modeling

Linear regression models were used, including baseline and multiple regression models.

Model performance was assessed using R-squared (R^2) and Root Mean Squared Error (RMSE).

Visualizations, including scatter plots and residual plots, were employed for model evaluation.

Models aimed to predict house prices based on various features, including bedrooms, bathrooms, condition, and more.

The project aimed to gain insights into the housing market through these models and metrics.



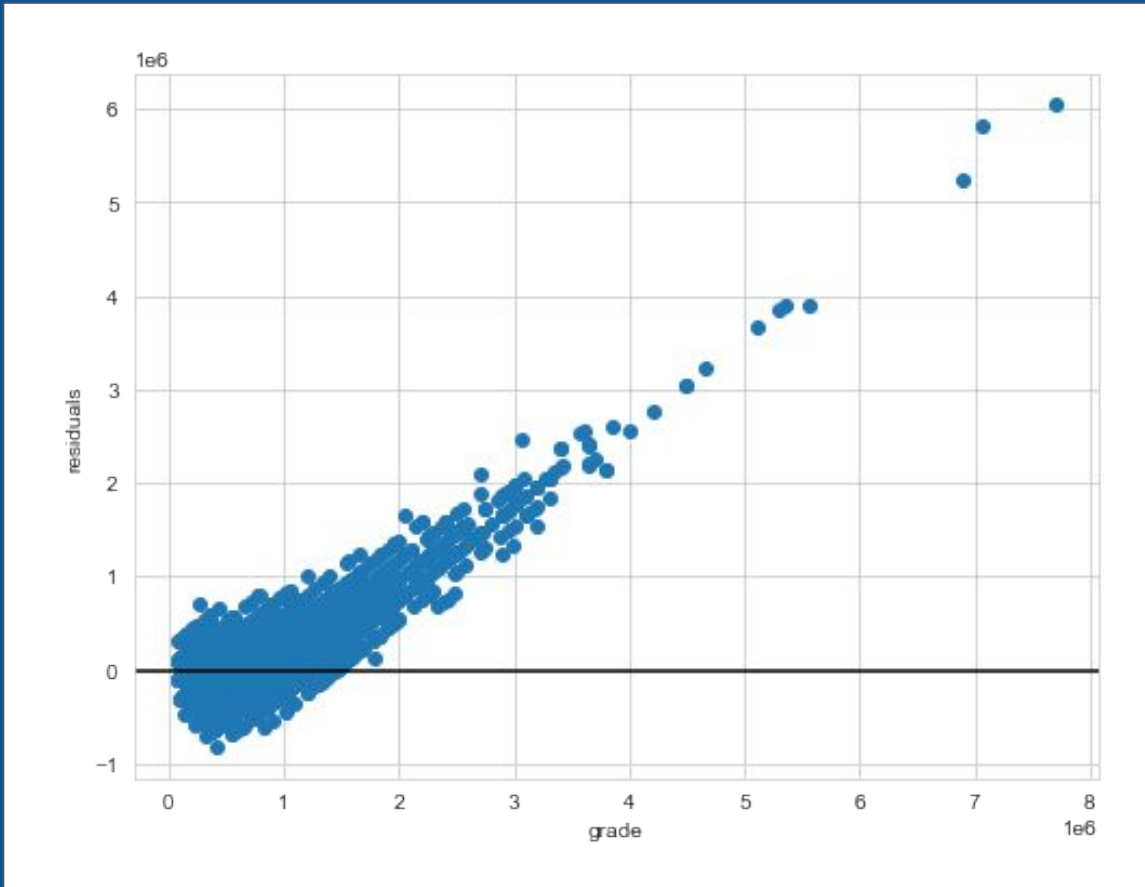
Regression Results

Key Factors: "bedrooms", "bathrooms", "sqft_lot", "sqft_basement", "sqft_above", "floors" are pivotal in determining house prices in the northwestern county.

Model Efficacy: Our top model captures 84.3% of price variations. This indicates that approximately 84.4% of the variance in housing prices is explained by the model. The chosen features (bedrooms, bathrooms, sqft_lot, sqft_basement, sqft_above, floors) are highly influential in predicting prices.

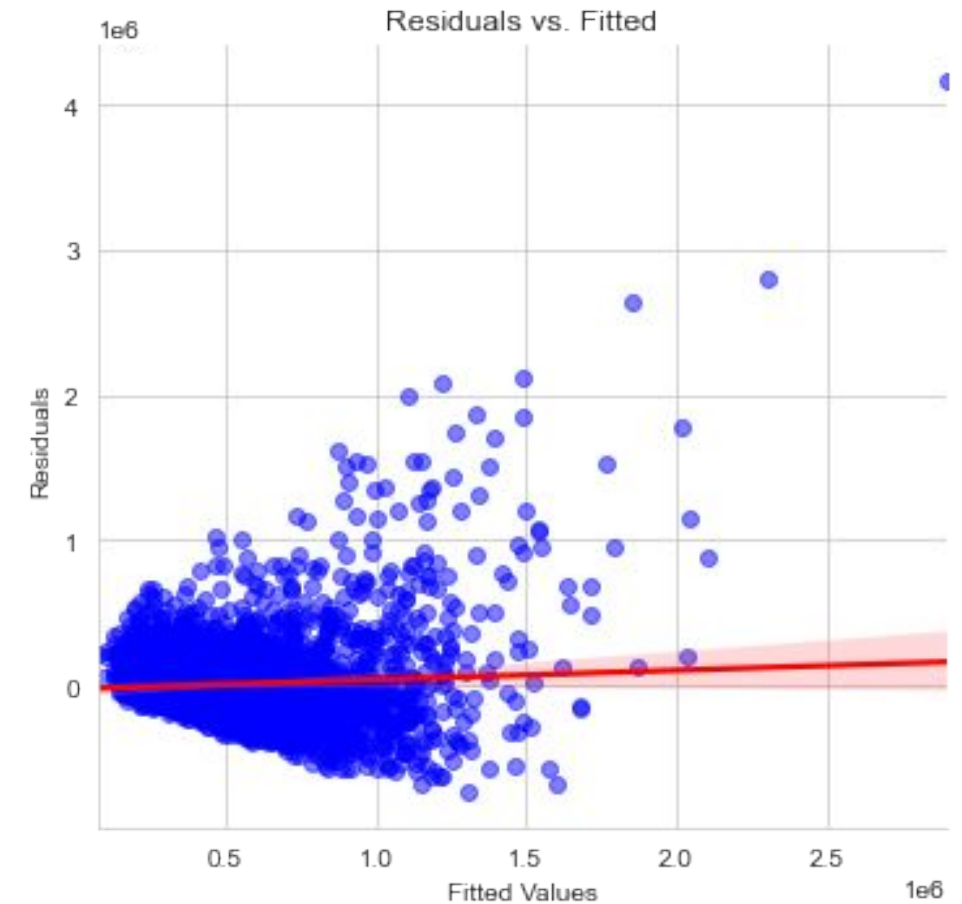


Baseline Model



The scatterplot shows negative values indicating that the model underestimated the prices, while positive values indicate overestimation.

Predictive Model



The scatterplot above shows that the model tends to make accurate predictions, with a majority of residuals clustering around zero.

Conclusions

Qualitative Aspects: The condition, grade, and view of a property are pivotal factors influencing home prices. Renovations aimed at enhancing these aspects can yield significant increases in property values.

Quantitative Features: Key quantitative features, such as the number of bedrooms, bathrooms, and total square footage, exhibit strong correlations with home prices. These features are essential considerations for homeowners and investors.

Significant Predictors: The regression models revealed that the features that strongly influence house prices include the number of bedrooms, bathrooms, square footage, condition, grade, and view.

However, number of bedrooms, bathrooms, square footage of the lot, square footage of the basement, square footage of house apart from basement and Number of floors (levels) in house seemed to be the best in explaining the house prices

Limitations

Data Completeness: The dataset did not cover all the features that influence house prices. For instance, the location or neighborhood of a property, local amenities, public transport accessibility, crime rates, or school quality can also play significant roles in determining prices.

Data Quality: There might be inconsistencies in the dataset that could have affected the results.

Multicollinearity: Features like 'sqft_living', 'sqft_above', and 'grade' might be correlated with each other, leading to multicollinearity. While we addressed this by removing 'sqft_living', other multicollinear variables might still be present. This can cause instability in regression coefficients, making interpretations difficult.

Time Sensitivity: House prices change over time due to economic, political, or environmental factors. The dataset did account for temporal aspects or trends in the housing market.

Recommendations

Client Guidance: Weichert Realtors should provide clients with personalized guidance on home renovations and improvements. Explain how enhancing qualitative aspects like condition, grade and view can boost property values, and suggest specific renovations aligned with their goals.

Educational Resources: Develop educational resources and materials emphasizing the importance of specific renovations, such as increasing number of bedrooms, bathrooms, square footage of the lot, square footage of the basement, square footage of house apart from basement and Number of floors in house. Optimizing these features can lead to higher estimated home values.

Maximizing ROI with the right Renovation Strategies: Emphasize on the importance of quality. Ensure customers understand that renovations should be done to a high standard of quality. Quality renovations not only increase property value but also attract potential buyers or renters.

Marketing Strategies: Highlight unique property features, such as waterfront views or exceptional grades, in marketing materials. Showcasing these features contribute to higher property values.

Cont...

Further Data Exploration: Exploring additional variables, such as location-related features or neighborhood characteristics, could provide deeper insights into price determinants. Proximity to schools, shopping centers, public transportation, parks, and the overall desirability of the neighborhood can greatly impact the price.

Risk Management: It's essential to acknowledge the potential for multicollinearity in the data, which can affect model accuracy. We recommend further research to mitigate this issue and ensure reliable predictions.

Advanced Modeling: Continue to invest in advanced modeling techniques and data analysis to refine pricing strategies and enhance predictive accuracy by taking into consideration the Market trends, Seasonal Variations and Economic Factors . This will ensure that clients receive the most accurate and up-to-date information.

Client Feedback: Encourage clients to provide feedback on their experiences and preferences. Use this feedback to further tailor services and recommendations to meet their individual needs.

THANK YOU

