



# **PREDICTING HOUSE PRICES IN KING COUNTY**

## **A MULTIPLE LINEAR REGRESSION APPROACH**

# CONTENT

- 01 INTRODUCTION
- 02 PROBLEM STATEMENT
- 03 OBJECTIVES
- 04 ANALYSIS
- 05 MODELLING
- 06 CONCLUSION
- 07 RECOMMENDATION

# INTRODUCTION

King County boasts of a robust economy driven by tech giants like Amazon, Microsoft, and Boeing, which continually attract a large workforce, driving the County's population through the roof, and with it the demand for homes..

In this project, we leverage a Multiple Linear Regression model to gain insights into the dynamic real estate market of this thriving region.

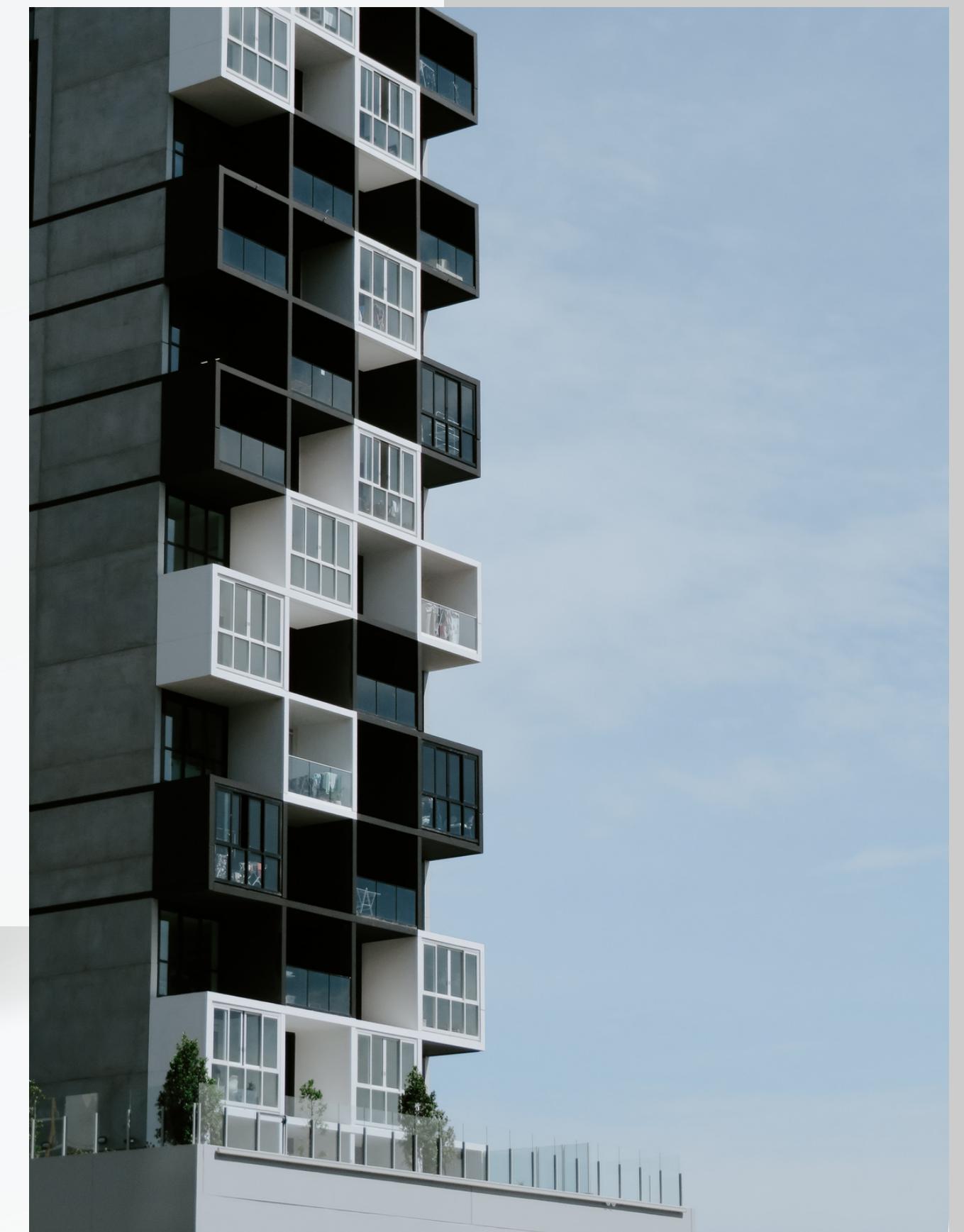
Our analysis is based on a comprehensive dataset from King County, which encompasses a multitude of factors influencing property prices.



# PROBLEM STATEMENT

The King County real estate market is highly competitive, with various stakeholders such as developers, online platforms, and established real estate companies vying for their share of the market.

In this dynamic real estate market, where economic conditions, housing demand, and external influences drive property prices, the importance of accurate pricing cannot be overstated.



*Photo by [Denise Jans](#) on [Unsplash](#)*

# OBJECTIVES

To explore and analyze the impact of numeric attributes on house prices in King County. Identify which numeric features have the most significant influence on pricing. Provide insights into how each unit increase or decrease in these attributes affects the final sale price. Generate recommendations for homeowners, buyers, and investors to optimize property attributes and investments based on numeric data.

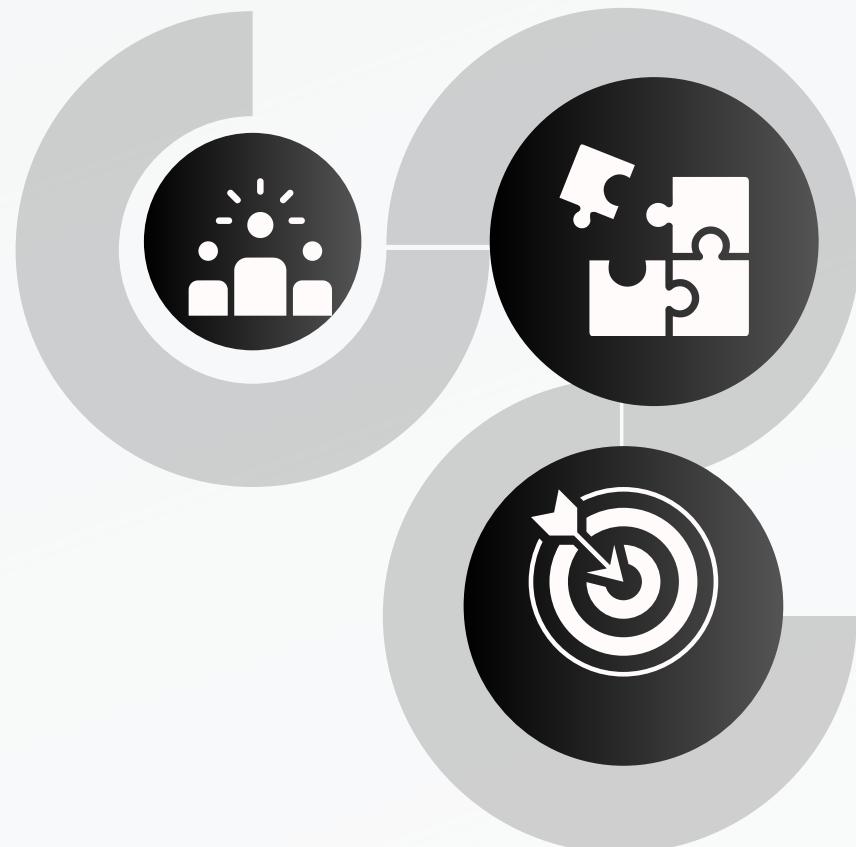
To investigate the influence of categorical attributes on house prices. Determine which categorical features, such as being on a waterfront or having a high-grade rating, command premium prices. Provide recommendations on how to leverage these categorical attributes to maximize property values. Assist stakeholders in making informed decisions based on categorical data.



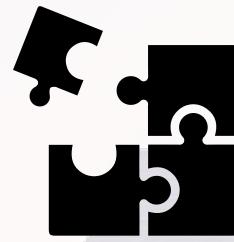
# OBJECTIVES

To create a precise property valuation model that calculates the cost of homes depending on a range of characteristics.

- 03** This model will utilize a property's characteristics, including but not limited to bedrooms, square footage, and more. By carefully selecting and incorporating these features, we intend to build a model that accurately reflects the diverse attributes influencing house prices.



# DATA CLEANING AND PREPARATION



## Handling Missing Values

Null values were replaced: 'waterfront' and 'view' were set to "unknown," while 'yr\_renovated' was imputed with the most common value, "0."



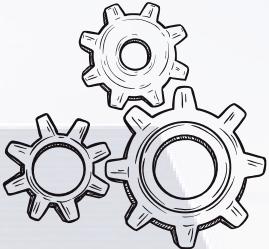
## Duplicate Entries

177 entries were flagged as duplicates. Further investigation revealed that some houses were sold multiple times at different prices and times.



## Handling Outliers

Outliers were retained because they represented genuine property attributes with valuable pricing information.



## Feature Engineering

New columns 'house age level' and 'seasons' were created from the existing columns. 'House age' was created from year built while seasons was created from date sold.

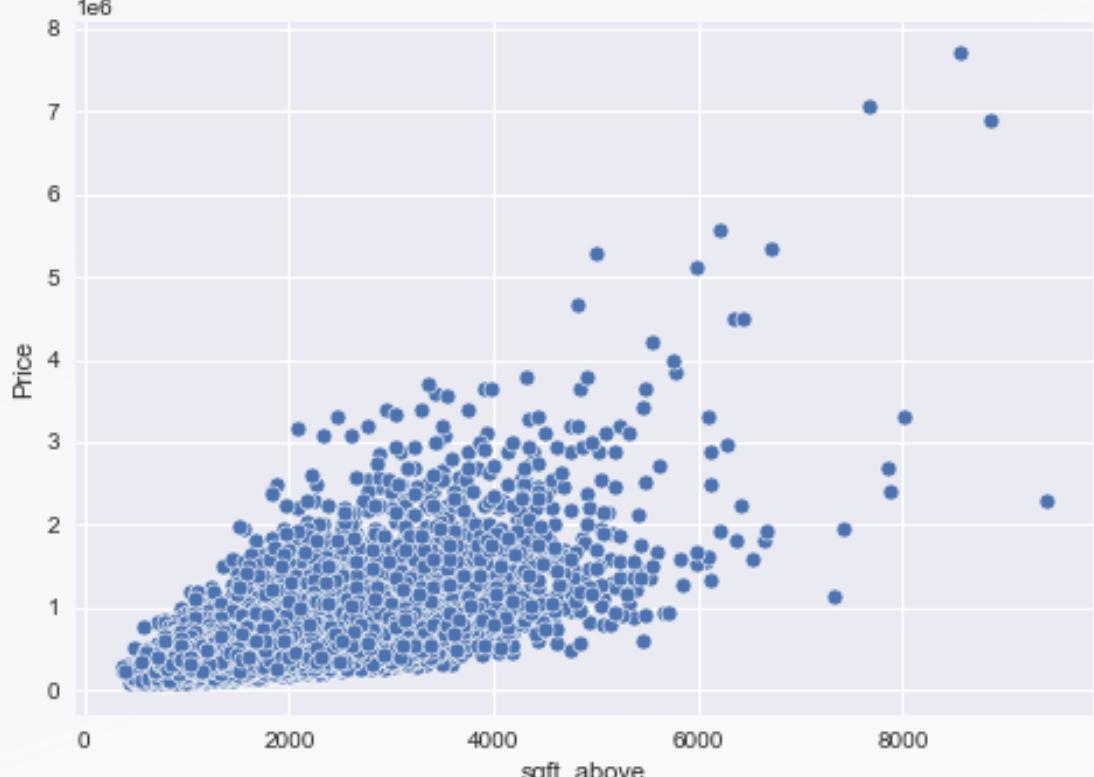
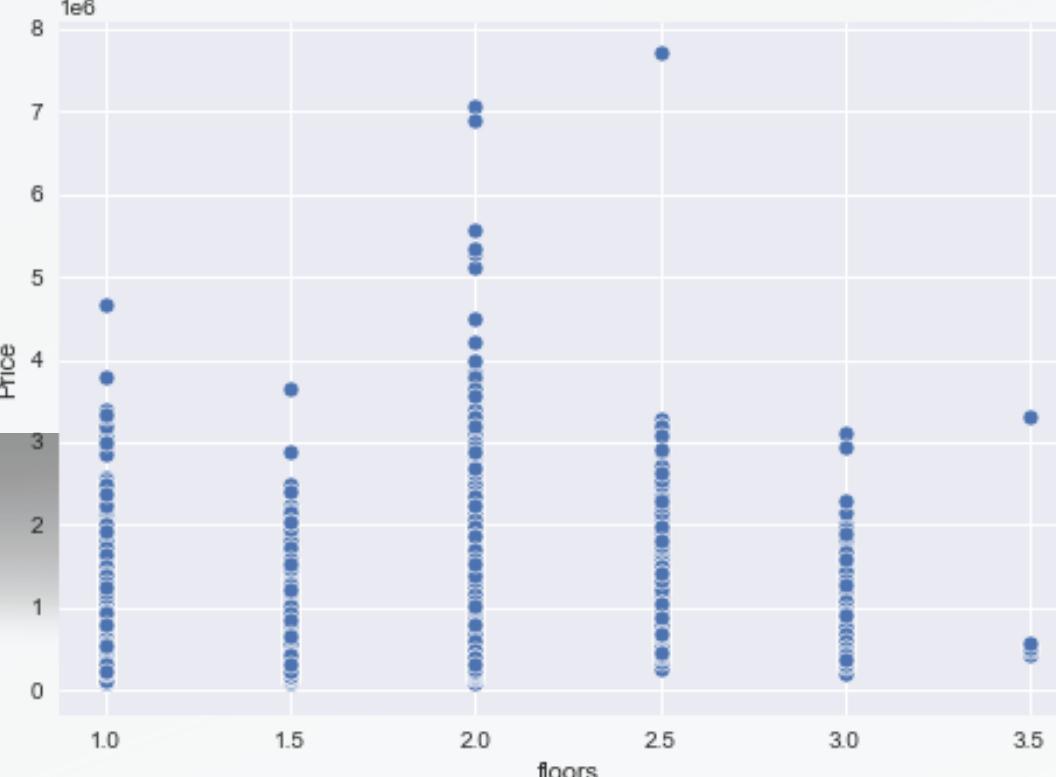
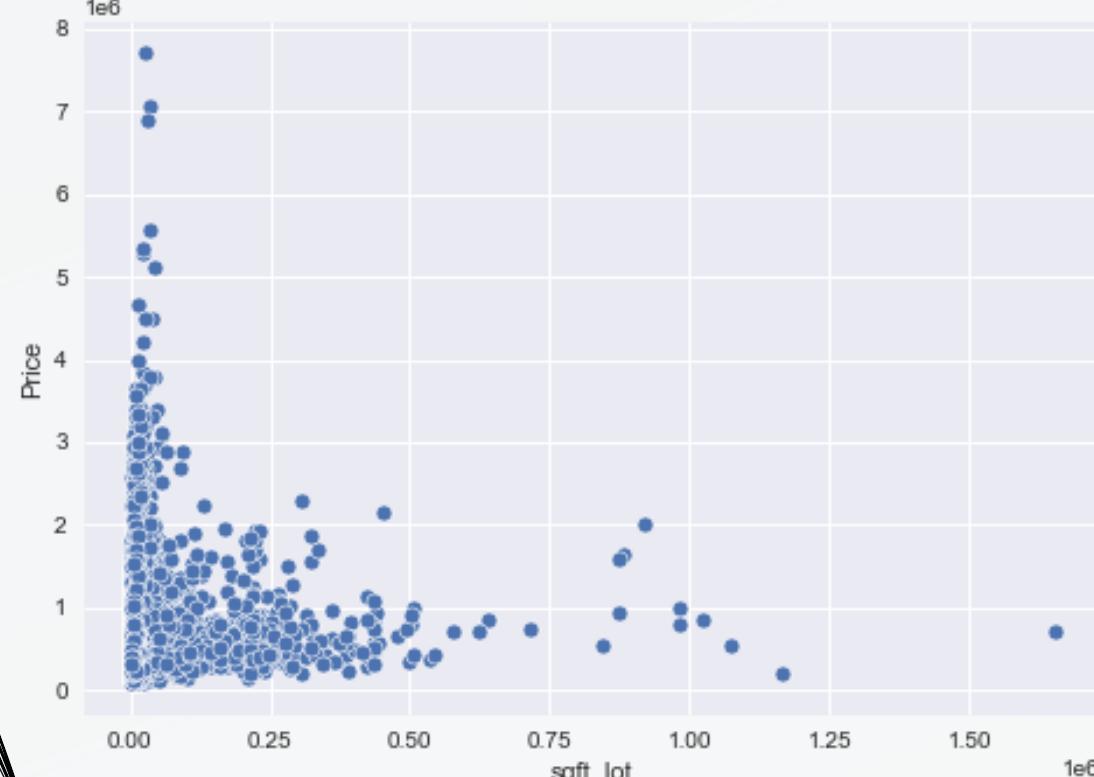
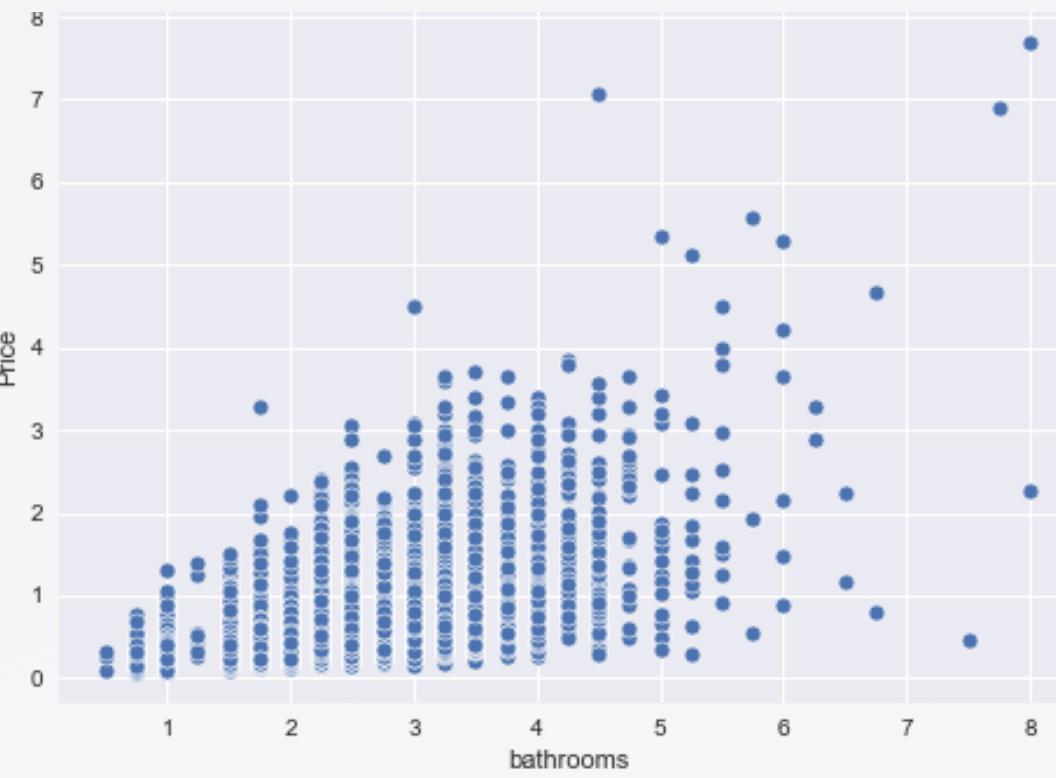
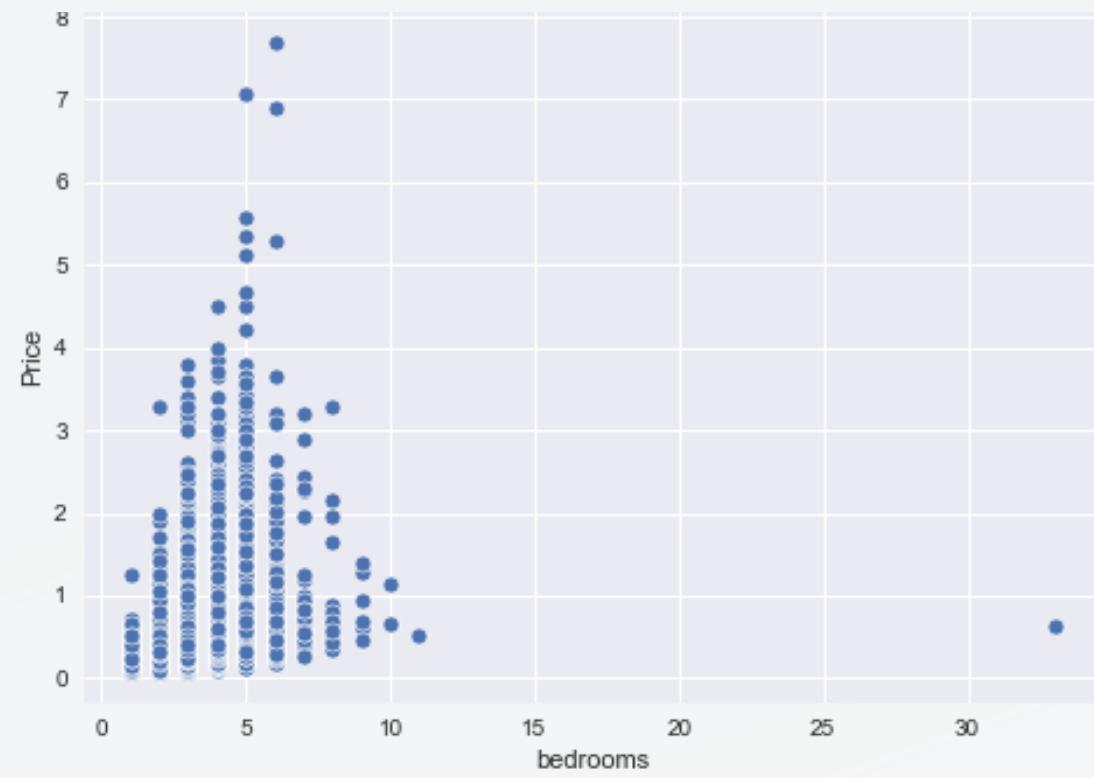
# ANALYSIS

Scatter plots were utilized to examine the relationships between the *price* (*target variable*) and various attributes such as bathrooms' '*sqft\_living*, *sqft\_above*, *sqft\_basement*, *bedroom*, *sqft\_lot*, and *floors*.

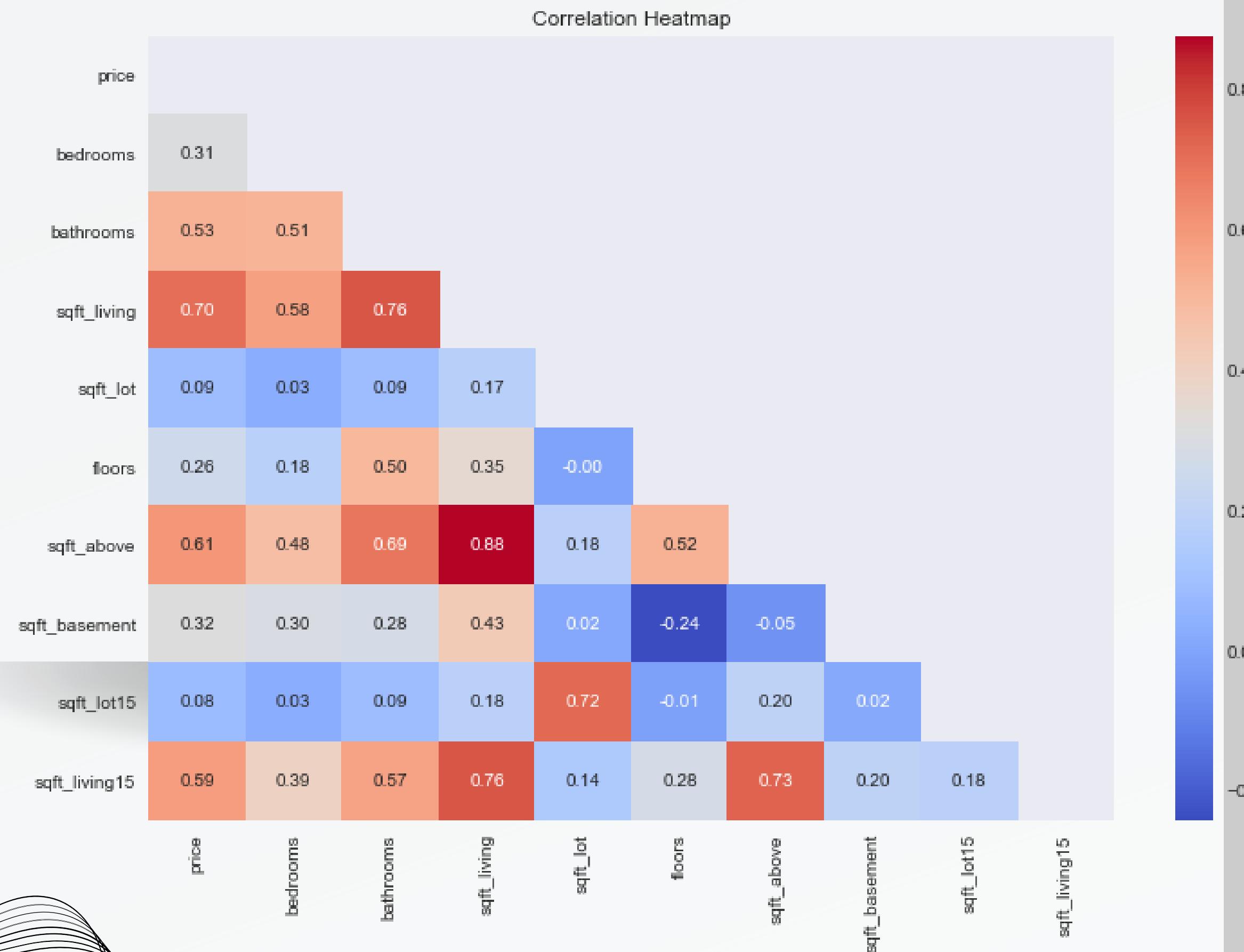
A heatmap was utilized to determine which feature exhibits the highest correlation with the target variable, *price*. The analysis revealed that the *sqft\_living* column possesses the most substantial correlation of 0.70 with price, indicating that it is a strong candidate to be the independent variable in the baseline model.



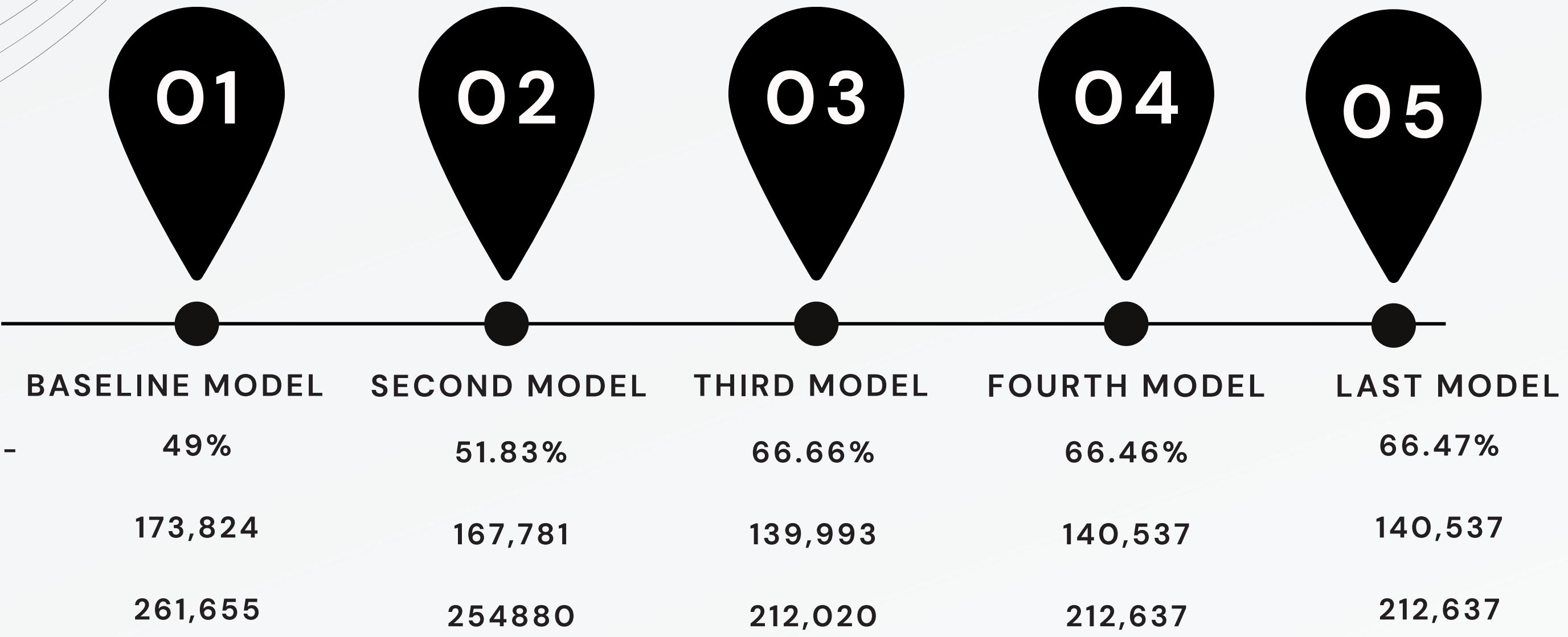
# ANALYSIS



# ANALYSIS



# MODELLING RESULTS



# CONCLUSION

**Waterfront:** Homes with a waterfront view have the most substantial positive impact on their prices. This feature is highly desirable and commands a premium.

**House Grade:** The quality and grade of the house play a crucial role in determining its price. Properties with higher grades, such as "Mansion" and "Luxury," have significantly higher values. Quality construction and design are valued by buyers.

**Square Footage:** More living space, including basements, has a positive effect on house prices. Larger homes tend to command higher values in the real estate market.

**Bathrooms and Floors:** Additional bathrooms and floors in a house contribute positively to its price. These features offer convenience and comfort, which are reflected in the property's value.

# CONCLUSION

**Lot Size:** Surprisingly, larger lot sizes, especially Lot 15, have a negative impact on prices. This suggests that smaller, more manageable lots are preferred and can even result in higher property values.

**House Age:** Older houses tend to be more expensive than newer ones. This might be due to historical or architectural significance associated with older properties. Buyers are willing to pay a premium for such houses.

**Bedrooms:** An increase in the number of bedrooms is associated with lower house prices. This finding may reflect buyer preferences, as larger homes with more bedrooms might cater to a different market segment.

# RECOMMENDATIONS

**Prioritize Waterfront Properties:** If you're a buyer looking for an investment or a dream home, consider waterfront properties. These offer not only a beautiful living environment but also a significant potential for property value appreciation.

**Enhance House Quality:** As a seller, focus on improving the quality of your property.

Consider renovations or upgrades to increase the house grade, which will positively impact your selling price.

Emphasize any unique design features or high-quality construction.



# RECOMMENDATIONS

**Highlight Square Footage:** When selling a property, make sure to highlight the square footage of the living space, including any basements. Potential buyers often place great importance on having enough space for their needs.

**Consider Additional Bathrooms and Floors:** If you're planning to invest in a property, consider adding more bathrooms or additional floors to enhance its value. These features are attractive to many buyers and can result in a higher selling price.



# RECOMMENDATIONS

**Optimize Lot Sizes:** If you have a property with a larger lot size, consider the possibility of subdividing it into smaller, more manageable lots. Smaller lots, especially those similar in size to Lot 15, appear to be preferred in the market and can potentially lead to better prices.

**Value Older Homes:** Older homes often come with historical and architectural charm. Sellers can emphasize these unique features to attract buyers who appreciate the character and history of older properties.



# RECOMMENDATIONS

**Optimize Bedroom Layouts:** Be mindful of the number of bedrooms in a property. Consider how they are laid out and whether the layout is appealing. Effective bedroom design can help maintain property appeal and value.



# LIMITATIONS

The analysis relies on available data, potentially missing critical variables that could influence real estate pricing.

DATA CONSTRAINTS

Economic shifts and government policies were excluded, which can have a substantial impact on the real estate market.

EXTERNAL VARIABLES

The model assumes linear relationships, neglecting potential nonlinear interactions that could exist in the data.

SIMPLIFIED MODEL

# NEXT STEPS

**By implementing these recommendations, stakeholders in the King County real estate market can further refine their models and decision-making processes, ultimately improving their ability to navigate this dynamic and competitive environment.**

## INCORPORATE ECONOMIC INDICATORS

To enhance the accuracy of market trend predictions, consider integrating economic indicators into the model. Variables such as local employment rates, income levels, and interest rates can offer valuable insights into housing market dynamics.

## ADVANCED PREDICTIVE MODELS

While the linear regression model has provided valuable insights, consider exploring advanced machine learning techniques such as gradient boosting and neural networks. These models can handle complex relationships and interactions within the data, potentially leading to more precise price forecasts.

# **PRESENTERS**

**Joy Ongutu**

**Grace Mutuku**

**Shuaib Mahamud**

**Kelvin Rotich**

**Peter Otieno**



**QUESTIONS?**

