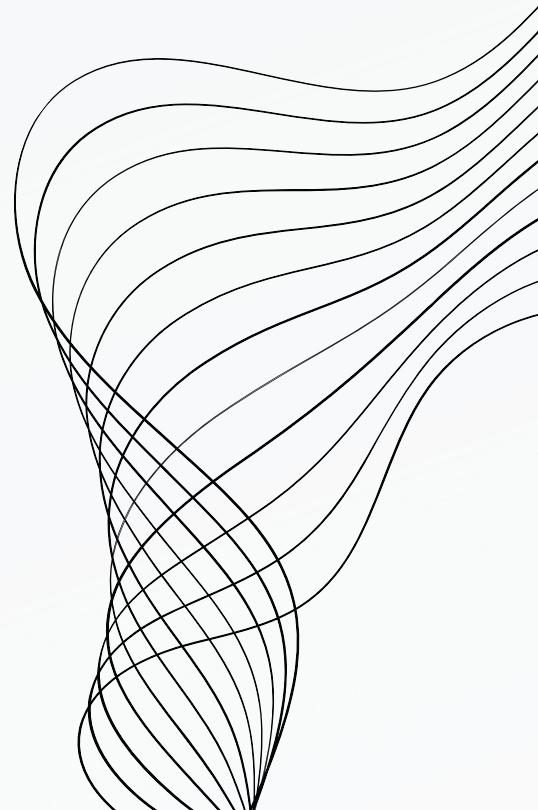




# KING COUNTY REAL ESTATE: CRAFTING COMPETITIVE PRICING STRATEGIES



# CONTENT

- 01** PROJECT OVERVIEW
- 02** BUSINESS PROBLEM & UNDERSTANDING
- 03** OBJECTIVES
- 04** DESCRIPTIVE ANALYSIS
- 05** INFERENTIAL STATISTICS
- 06** REGRESSION ANALYSIS
- 07** RECOMMENDATIONS

# PROJECT OVERVIEW

This project aims to assist a commercial real estate agency in King County by developing a data-driven pricing strategy. Using the King County House Sales dataset, we'll identify key factors affecting house prices, analyze location's impact on valuations, and create a regression model to understand the factors that affect house prices.





# BUSINESS UNDERSTANDING

The real estate agency aims to develop a pricing strategy by understanding the key factors influencing house prices. To achieve this, we're tasked with identifying the main drivers of house prices. The client desires accurate insights based on comprehensive data from the King County real estate market, encompassing historical sales, and other relevant features.

# OBJECTIVES



What are the  
key predictors  
of house prices?



How does  
location affect  
house prices  
within the  
region?



Can we predict  
the price of a  
house based on  
its attributes?

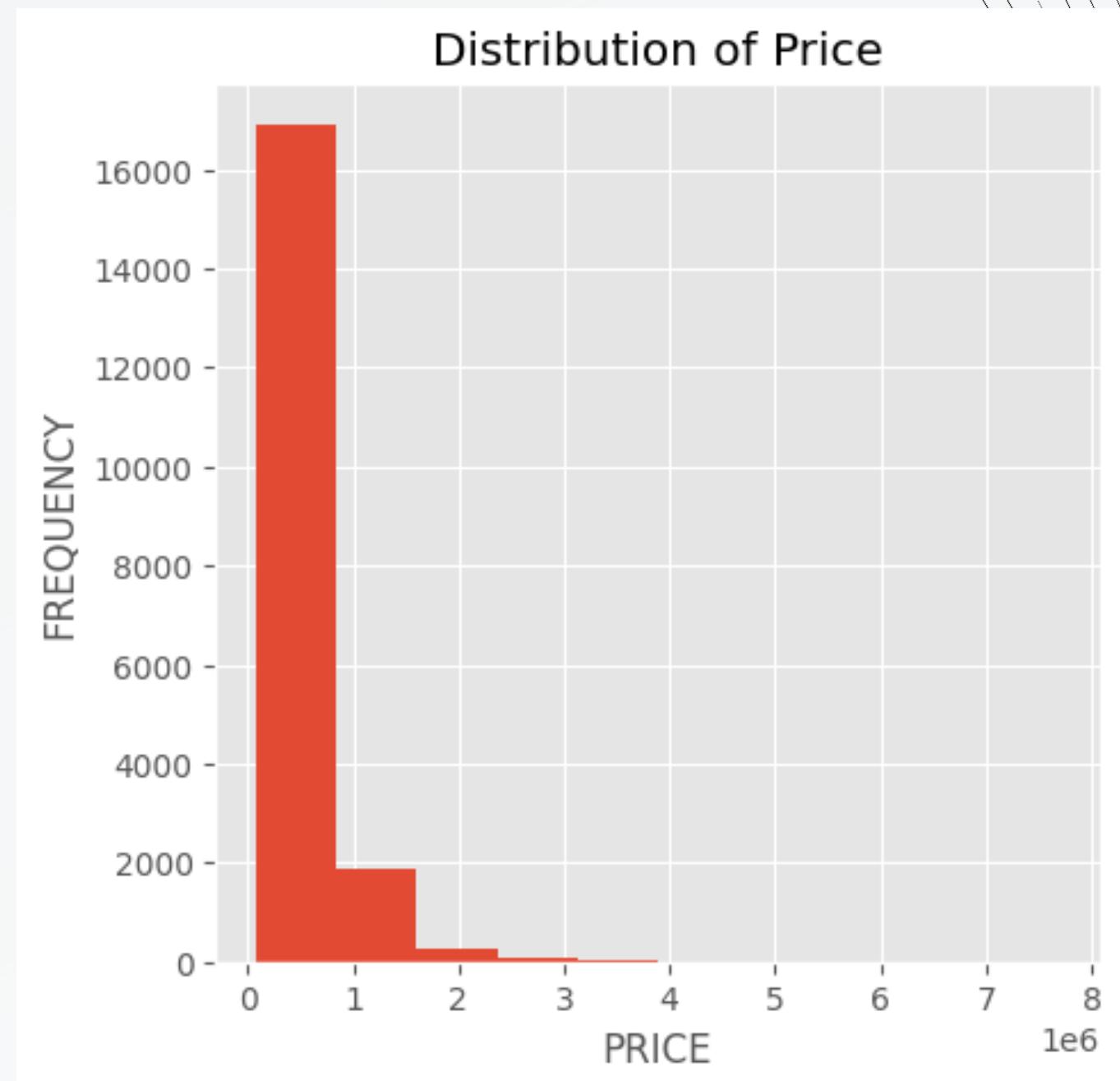
# DESCRIPTIVE ANALYSIS

In this section, we delve into key insights derived from our descriptive analysis of housing data. Through comprehensive visualizations and statistical summaries, we uncover crucial factors influencing house prices. Our exploration aims to provide valuable insights to guide strategic decisions for the real estate company.



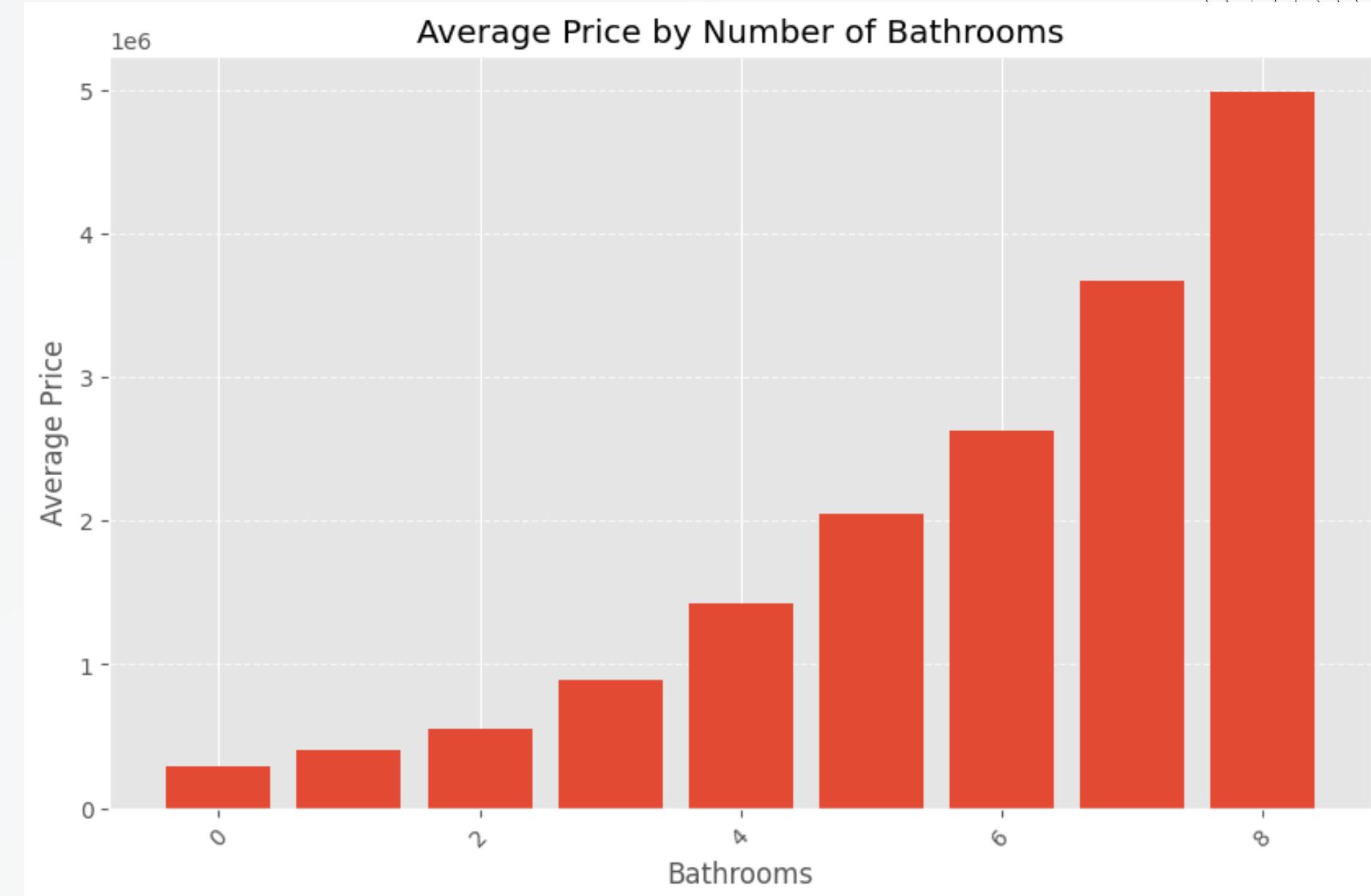
# PRICE

- House prices in Kings County span from \$78,000 to \$7,700,000, showcasing a broad spectrum of affordability within the region.
- The median price of \$450,000 serves as a central point of reference, indicating that half of the houses in Kings County are priced below this value and half are priced above it.
- The right-skewed distribution suggests that the majority of houses in Kings County are priced at the lower end of the spectrum, with fewer houses commanding higher prices. This skewness implies a prevalence of more affordable housing options relative to luxury properties within the county.



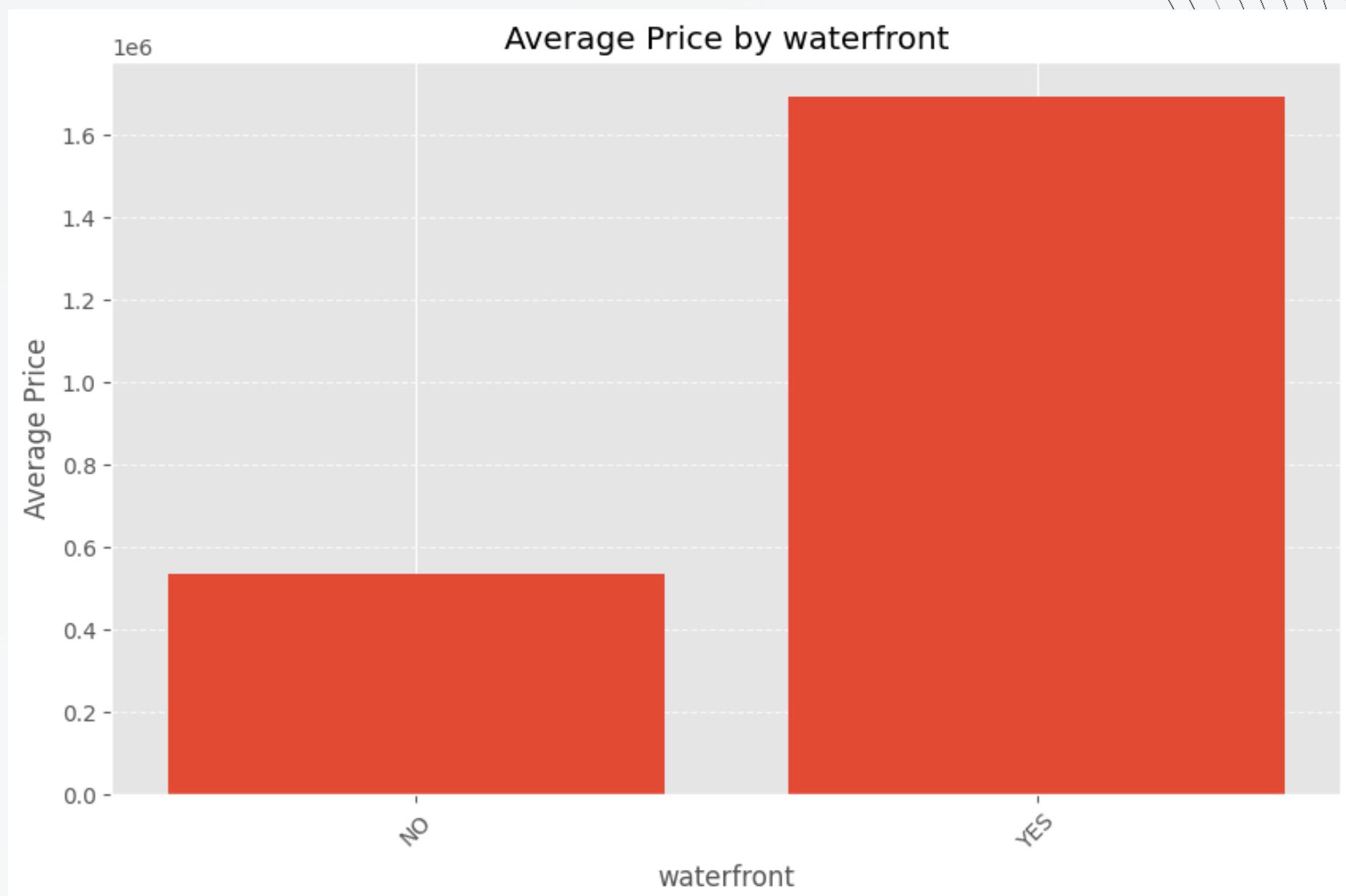
# BATHROOMS

- Visualizations demonstrate a consistent trend where houses with a greater number of bathrooms exhibit higher average prices.
- This trend holds true across the spectrum of bathrooms, ranging from 0 to 8, indicating a premium associated with increased bathroom count for potential buyers.



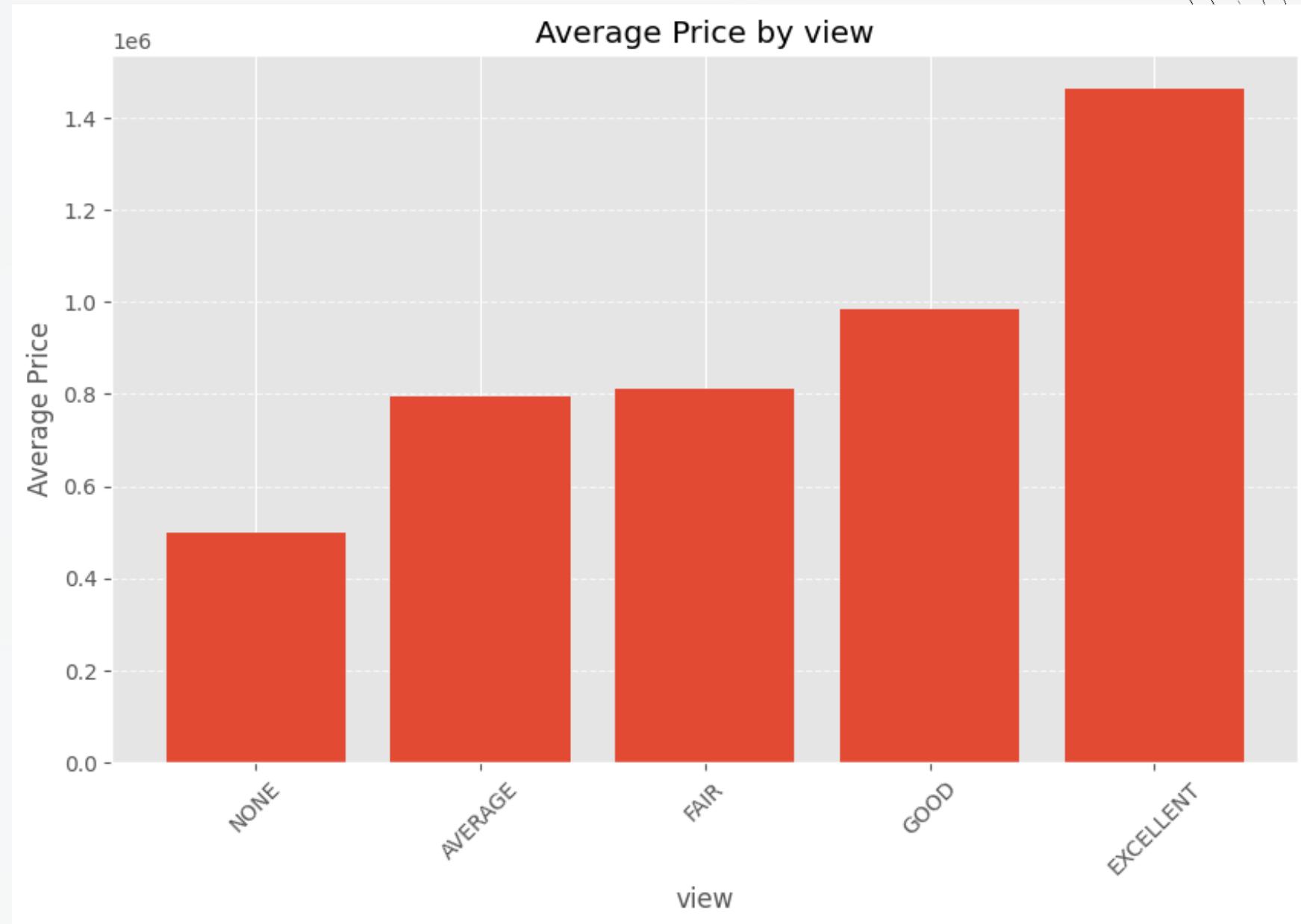
# WATERFRONT

- Analysis confirms a substantial price differential between houses with waterfront views (labeled "Yes") and those without (labeled "No").
- The waterfront designation significantly influences pricing, reflecting the desirability of properties offering scenic water views.



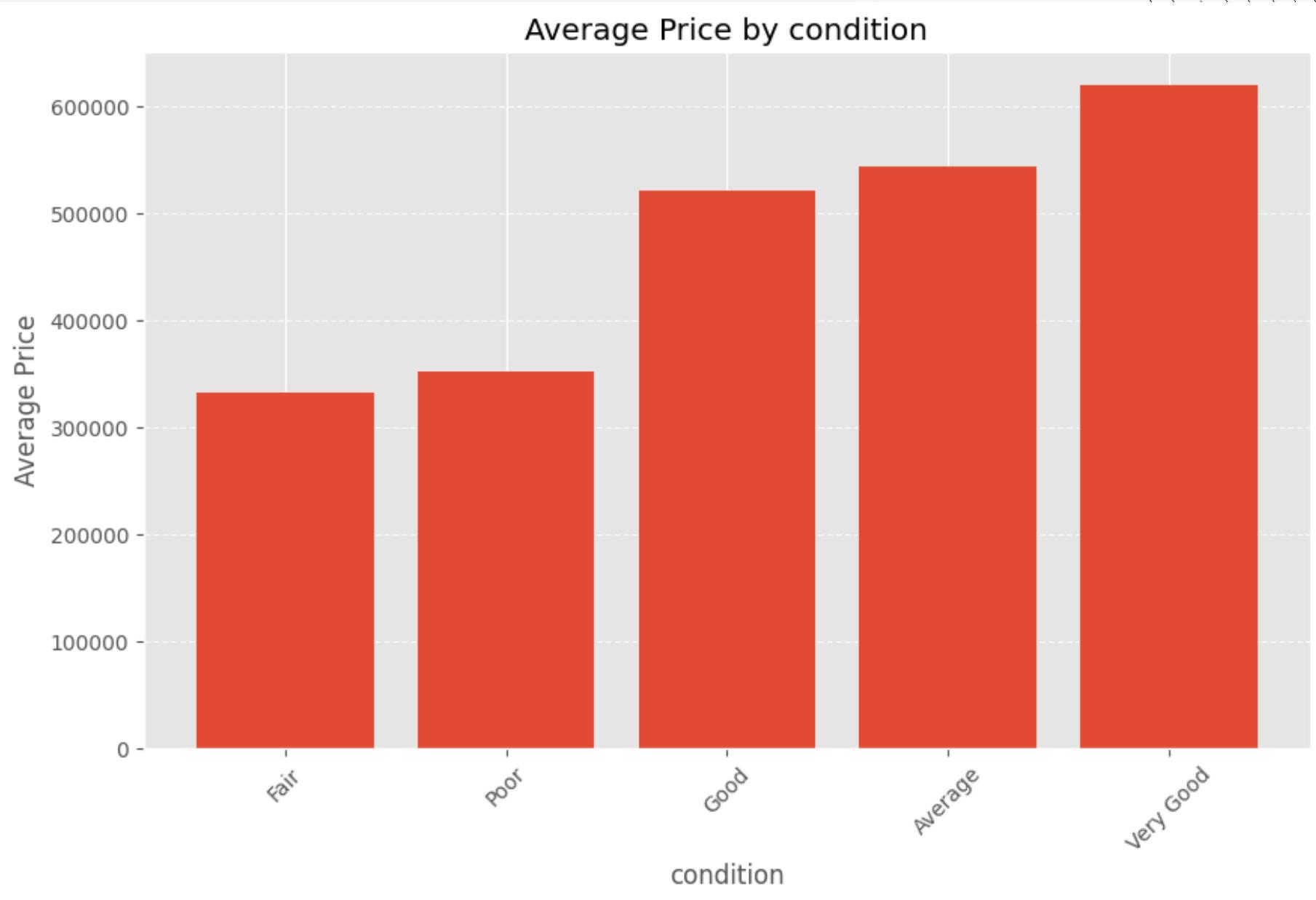
# VIEW

- Houses categorized with view qualities of Good and Excellent demonstrate notably higher average prices compared to those with None, Average, or Fair views.
- This highlights the impact of scenic views on property valuation, with superior views correlating with higher prices.



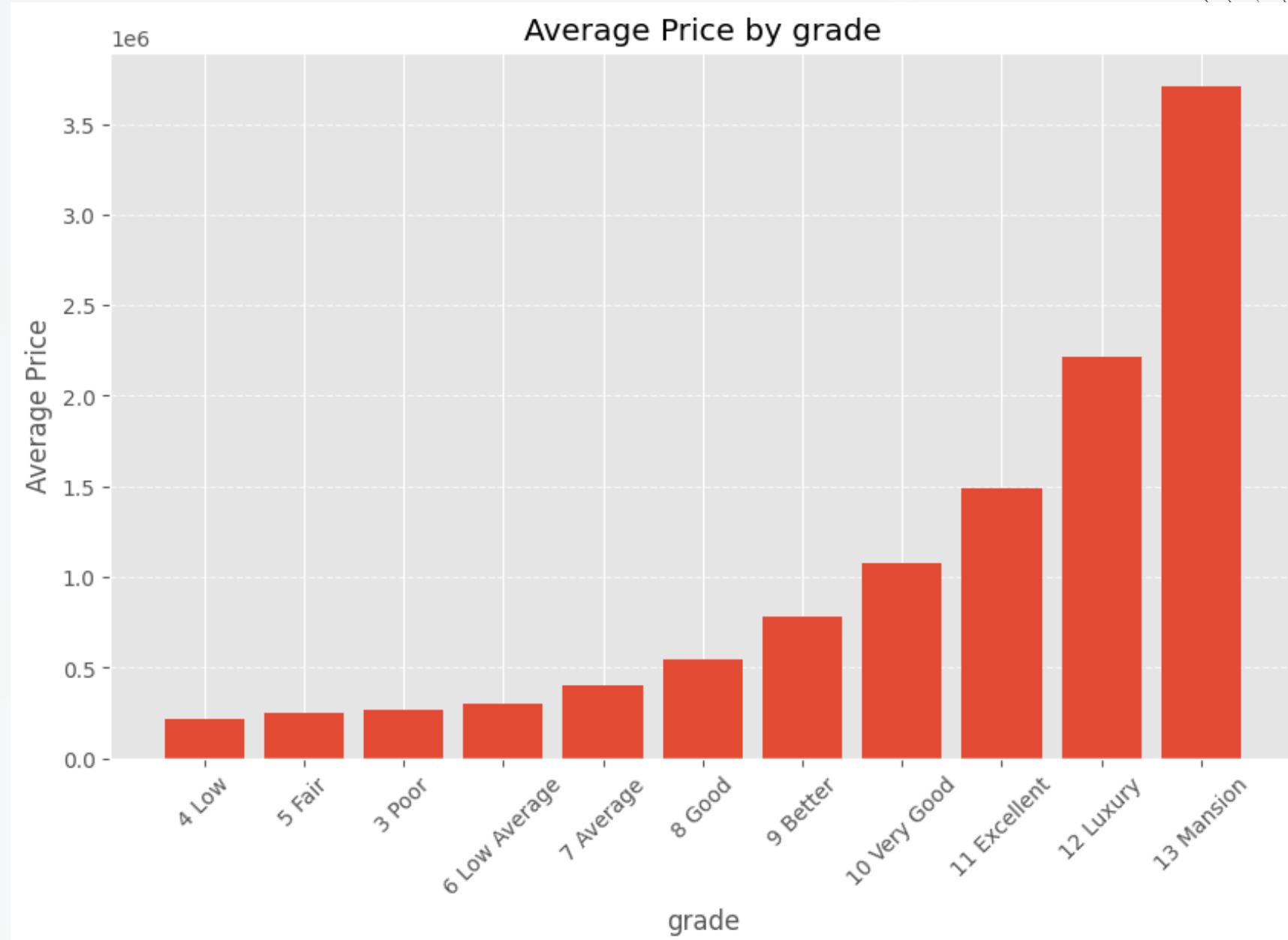
# CONDITION

- Properties labeled with conditions of Good, Average, and Excellent command higher average prices relative to those with lower condition ratings such as Poor or Fair.
- The condition rating serves as a crucial determinant of property value, reflecting the perceived quality and upkeep of the house.



# GRADE

- Houses grade as Excellent, Luxury, and Mansion are associated with progressively higher average prices, indicative of premium features and construction quality.
- Grade serves as a key indicator of the overall caliber and luxury level of the property, influencing buyer preferences and pricing.



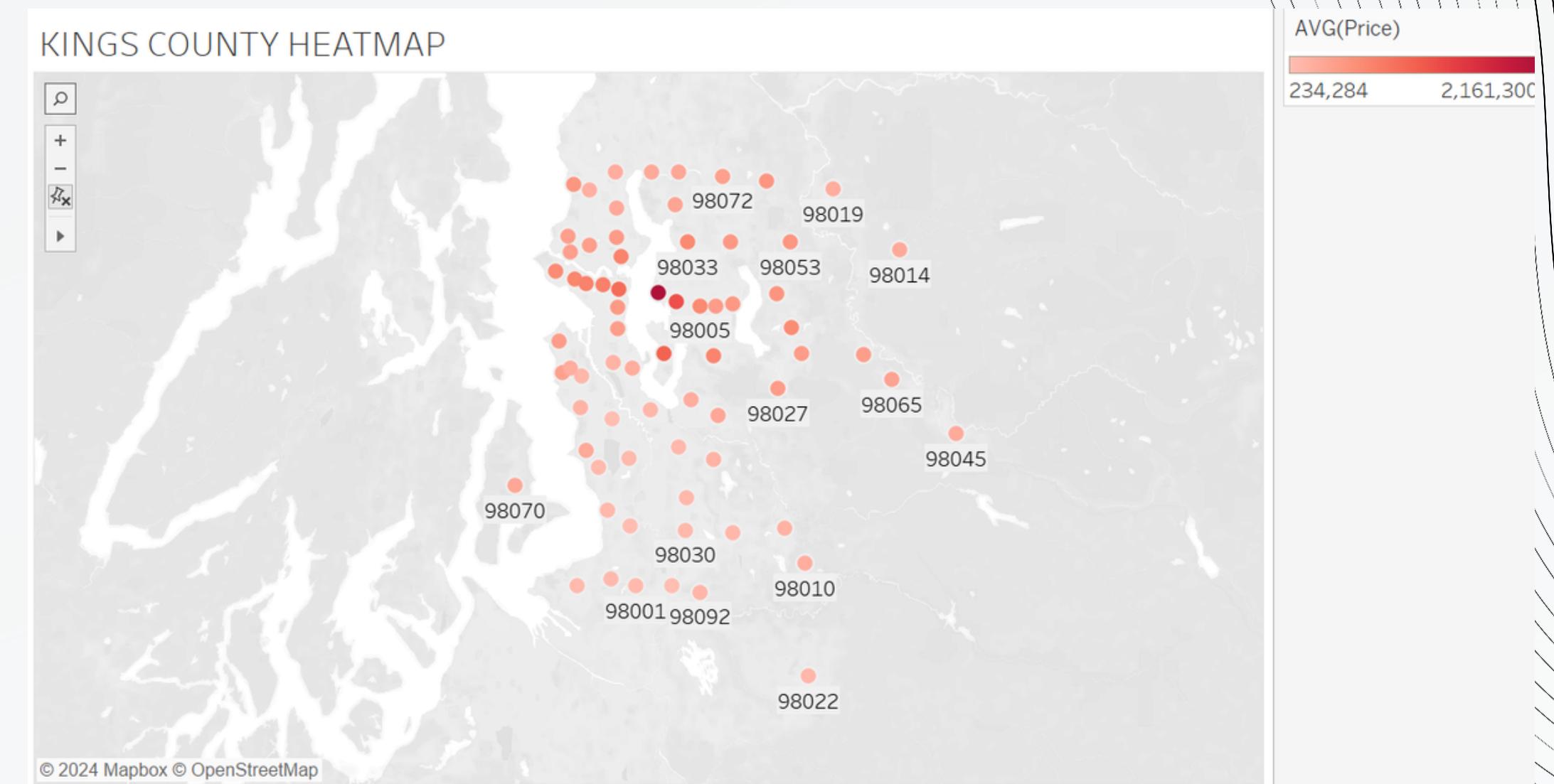
# BEDROOMS

- The analysis reveals that the average price of houses tends to increase with the number of bedrooms, reaching a peak at 8 bedrooms. Beyond this threshold, the average price begins to decrease, suggesting a diminishing return on additional bedrooms.



# HEATMAP

- Heatmap Analysis Reveals highest-priced houses concentrated in central region (darker shades of red).
- This is due to the proximity to amenities, historical significance, architectural charm, limited land availability.



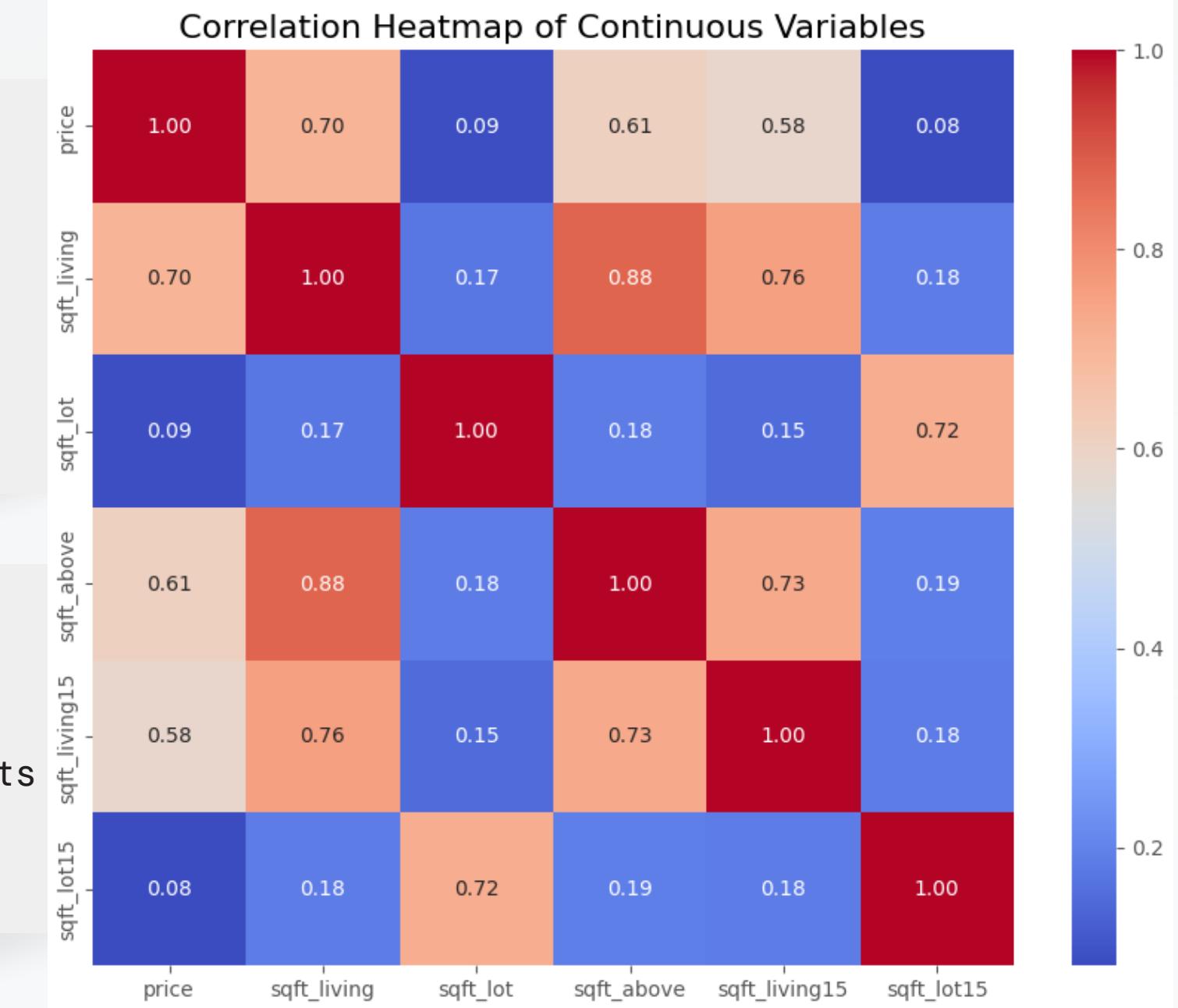
# INFERENCE STATISTICS INSIGHTS



A strong positive linear correlation exists between house prices and variables representing living space (`sqft_living`, `sqft_above`, `sqft_living15`), indicating that larger living areas correspond to higher prices.



However, our correlation heatmap reveals notable multicollinearity among certain pairs of variables: `sqft_living` and `sqft_living15`, `sqft_living` and `sqft_above`, and `sqft_lot` and `sqft_lot15`. This suggests redundancy in our regression analysis and warrants careful consideration during model development.



# REGRESSION ANALYSIS

In our regression analysis, we explored the relationship between various predictor variables and house prices. We developed three models: a simple linear regression model, a multiple regression model with two features, and another multiple regression model with many features. Through these analyses, we aimed to understand the factors influencing house prices and to develop pricing strategies.



# MODEL ANALYSIS

## MODEL 1

- R-squared: 0.490
- Adjusted R-squared: 0.489
- F-statistic: 1.470e+04 (p-value close to 0)
- Coefficient of sqft\_living: \$277.5570 (increase in price per sqft)
- Approximately 49% of the variance in house prices can be explained by the square footage of living space.

## MODEL 2

- R-squared: 0.507
- F-statistic: 7886 (p-value close to 0)
- Coefficient of sqft\_living: \$316.4133 (increase in price per sqft)
- Approximately 51% of the variance in house prices can be explained by square footage of living space and the number of bedrooms.

## MODEL 3

- R-squared: 0.581
- F-statistic: 2660 (p-value close to 0)
- Coefficient of sqft\_living: \$248.2605 (increase in price per sqft)
- Coefficient of bedrooms: -\$48,620 (decrease in price per bedroom)
- Approximately 58.1% of the variance in house prices can be explained by features including square footage of living space, number of bedrooms, and other factors.

# RECOMMENDATIONS



Incorporate identified factors into pricing strategy: Adjust listing prices based on factors like number of bathrooms, presence of waterfront views, and overall condition of the house to align with market demand and maximize profitability.



Utilize insights for targeted marketing: Tailor marketing efforts towards properties with desirable features such as waterfront views or higher-grade ratings to attract potential buyers and optimize sales.



Consider local market dynamics: Recognize the significance of location, as evidenced by the best-priced houses being in the central region of Kings County. Utilize this insight to focus resources on areas with high demand or potential for growth to maximize returns.

# NEXT STEPS

Develop advanced predictive models to forecast future house prices based on historical data and key predictors identified in the analysis. These models will provide valuable insights for pricing decisions and market forecasting.

## PREDICTIVE MODELING

Implement market segmentation techniques to identify distinct buyer segments based on preferences and demographics. Tailoring marketing strategies and pricing approaches to these segments will enhance targeting and effectiveness.

## MARKET SEGMENTATION

Conduct spatial analysis to identify spatial patterns and trends in house prices across different neighborhoods. Geographic visualization techniques will help pinpoint hotspots and areas of potential investment.

## SPATIAL ANALYSIS

# NEXT STEPS

Establish mechanisms for ongoing monitoring of the housing market, utilizing automated data collection and analytics tools. This proactive approach will enable timely decision-making and adaptation to changing market conditions.

## CONTINUOUS MONITORING

Create a feedback loop to gather insights from users of the analysis, such as real estate agents and decision-makers. Incorporating user feedback ensures that the analysis remains relevant and impactful.

## FEEDBACK LOOP

Enhance the dataset with additional relevant data sources, such as demographic and economic indicators. This will provide richer context and improve the accuracy and robustness of the analysis.

## DATA ENRICHMENT

# CONCLUSION

In conclusion, leveraging insights from our regression model, descriptive and inferential statistics analysis can significantly enhance decision-making in the real estate industry. By incorporating data-driven strategies into pricing and marketing efforts, we can effectively position properties in the market and capitalize on emerging trends.

# OUR TEAM

Winnie  
Osolo

Felix  
Mburu  
Njoroge

Violet  
Musyoka

Alex Korir

Lydia  
Masabarakiza

# THANK YOU

