



# HOUSE PRICES PREDICTION IN NORTHWESTERN COUNTY



# PROJECT OVERVIEW

---

- The dataset: The dataset consists of historical data on house prices and their corresponding features. It provides a rich source of information to analyze and build a predictive model.
- Data preprocessing: The dataset undergoes data cleaning and preprocessing steps, including handling missing values, and addressing any outliers or data inconsistencies.



# PROJECT OVERVIEW

---

- Exploratory Data Analysis (EDA): EDA is performed to gain insights into the relationships between the features and the target variable (house prices). Visualizations and statistical analyses are utilized to identify patterns, trends, and correlations in the data.



# BUSINESS UNDERSTANDING

---

- In this project, the aim is to analyze house sales in a northwestern county using multiple linear regression modeling. By leveraging historical data on house sales and associated factors, we seek to gain insights into the dynamics of the local real estate market.



# BUSINESS UNDERSTANDING

---

- This is to help Homeowners and Home Buyers:
  - Homeowners looking to sell their property can leverage the model's predictions to estimate the potential selling price and make informed decisions about pricing and negotiation.
  - Home buyers can use the predictions to assess whether a listed house is reasonably priced, helping them in their decision-making process.



# PROBLEM STATEMENT

---

- The objective is to address the challenges faced by various homeowners and home buyers in the housing market. Specifically, we aim to identify the key factors influencing house sale prices and develop a predictive model to accurately estimate the selling price of houses based on these factors.

# MAIN OBJECTIVE

---

- The main objective of this project is to provide actionable insights and decision-making support to homeowners and home buyers in the housing market by leveraging multiple linear regression modeling. By understanding the relationships between house characteristics (such as square footage, number of bedrooms, location, amenities) and sale prices, we aim to enhance pricing strategies, optimize property values, and facilitate informed negotiations.

# SPECIFIC OBJECTIVES

---

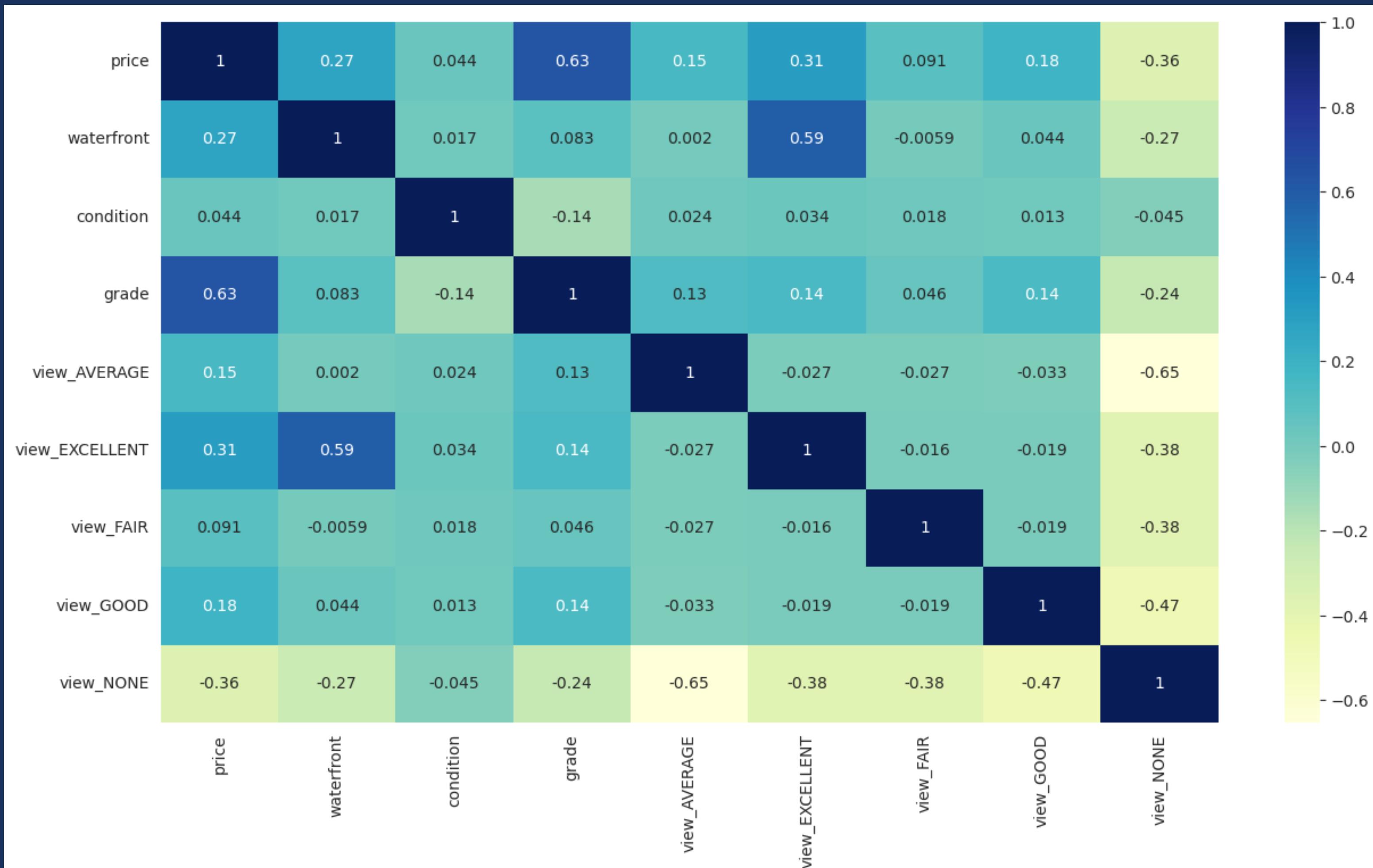
- Develop a model that accurately predicts house sale prices based on relevant features.
- Identify the most significant factors that influence house prices in the northwestern county, prioritising features that have the greatest impact.
- Provide homeowners and home buyers with a comprehensive understanding of the local housing market dynamics, enabling them to make data-driven decisions regarding house pricing.



# MODELING



# CORRELATION



## CORRELATION

- Waterfront: There is a moderate positive correlation between waterfront properties and price. This suggests that houses with a waterfront view tend to have higher prices.
- Grade: There is a strong positive correlation between the grade of the house and its price. Higher-grade houses tend to have higher prices, indicating that the quality and features of the house are important factors in determining its value.



## CORRELATION

- View: The presence or quality of the view is positively correlated with the price. Houses with better views (view\_EXCELLENT, view\_GOOD) tend to have higher prices, while houses with no view (view\_NONE) have a negative correlation with the price.
- Condition: The correlation between condition and price is very weak, indicating that the overall condition of the house may not have a significant impact on its price.



## REGRESSION RESULTS – MODEL PERFORMANCE

- **R-squared (uncentered):** 0.883

The model explains approximately 88.3% of the variance in the prices of the houses.

- **F-statistic:** 1.178e+04

Our model is statistically significant



## REGRESSION RESULTS - COEFFICIENTS

- Waterfront coeff: \$478,000

Holding all other variables constant, having a waterfront view this shows the associated average price increase.

- Grade: \$61,490

Holding all other variables constant, a one-unit increase in the grade is associated with an increase of \$61,490 in the price



## REGRESSION RESULTS – COEFFICIENTS

- View Excellent: \$101,700

Holding all other variables constant, having an excellent view rating is associated with an increase of \$101,700 in the price

- Sqft\_living: \$215.7328

Holding all other variables constant, a one-unit increase in the square footage of living space is associated with an increase of \$215.73 in the price.



# CONCLUSION

- The model's performance and significance make it a valuable tool for understanding the housing market.
- From the model Several predictor variables emerged as significant contributors to house prices. Notably, the presence of a waterfront, condition, grade, and different view categories (average, excellent, none) showed strong associations with house prices.
- Additionally, factors such as the number of bedrooms, bathrooms, square footage of living area, and the age of the house played significant roles in determining the price.



# RECOMMENDATIONS

- The model should be further refined and validated to ensure the model's reliability and applicability in real-world scenarios, moreover, further analysis can enhance its robustness and interoperability.
- The house prices either to sell or purchase should be determined by the location that's the accessibility of the waterfront, presence of a good view and a better or a higher grading.