

# KING COUNTY REAL ESTATE ANALYSIS

GROUP 3





# **Our Team**

**CRYSTAL WANJIRU**

**SCRUM**

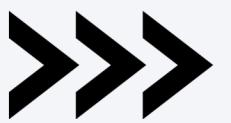
**RISANUEL OBONDO**

**TIMOTHY LENKU**

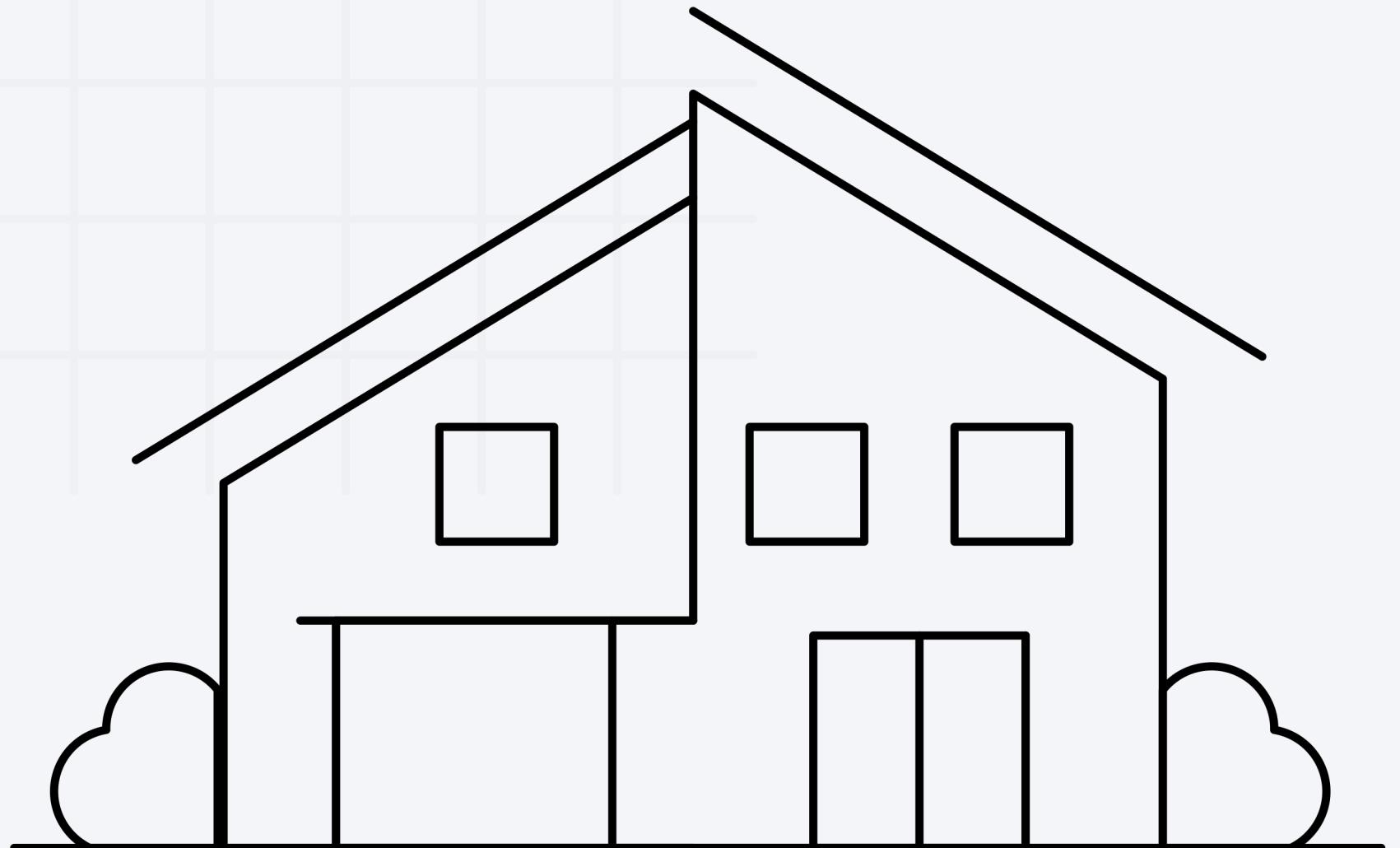
**WILSON MUTUNGU**

**LAUREEN CHEPKOECH**

**VICTORINE IMBUHILA**



# INTRODUCTION



King County, Washington, is a real estate powerhouse, pivotal to the region's economic growth. Influenced by location appeal, property features, and market trends, the diverse housing market in King County thrives due to a robust economy, tech industry presence, and its scenic beauty, offering a range of housing options, attracting both homebuyers and investors.



# PROBLEM STATEMENT

The booming King County real estate market lacks clear, actionable insights into the specific factors influencing property prices, hindering the ability of investors to make informed decisions.



Our aim is to offer or provide homeowners and real estate professionals with dependable insights for navigating the ever-changing King County real estate market with confidence.



# OBJECTIVES

01

To determine the main features that affect the value of a home. See how specific features relate to the housing price

03

To quantify the effect of the above features on the value of a home.

02

To check if there is a difference in the price of the renovated houses and those that were not renovated

04

To Develop a model to predict the home value.



# Data understanding

This project uses King County  
house sales dataset



## Data Composition

This data contains information about the houses in the northwestern county. This is found in kc\_house\_data.csv.

## Data Analysis

We use Pandas and NumPy for data cleaning and analysis, Tableau, seaborn and Matplotlib for data visualization

## Modelling

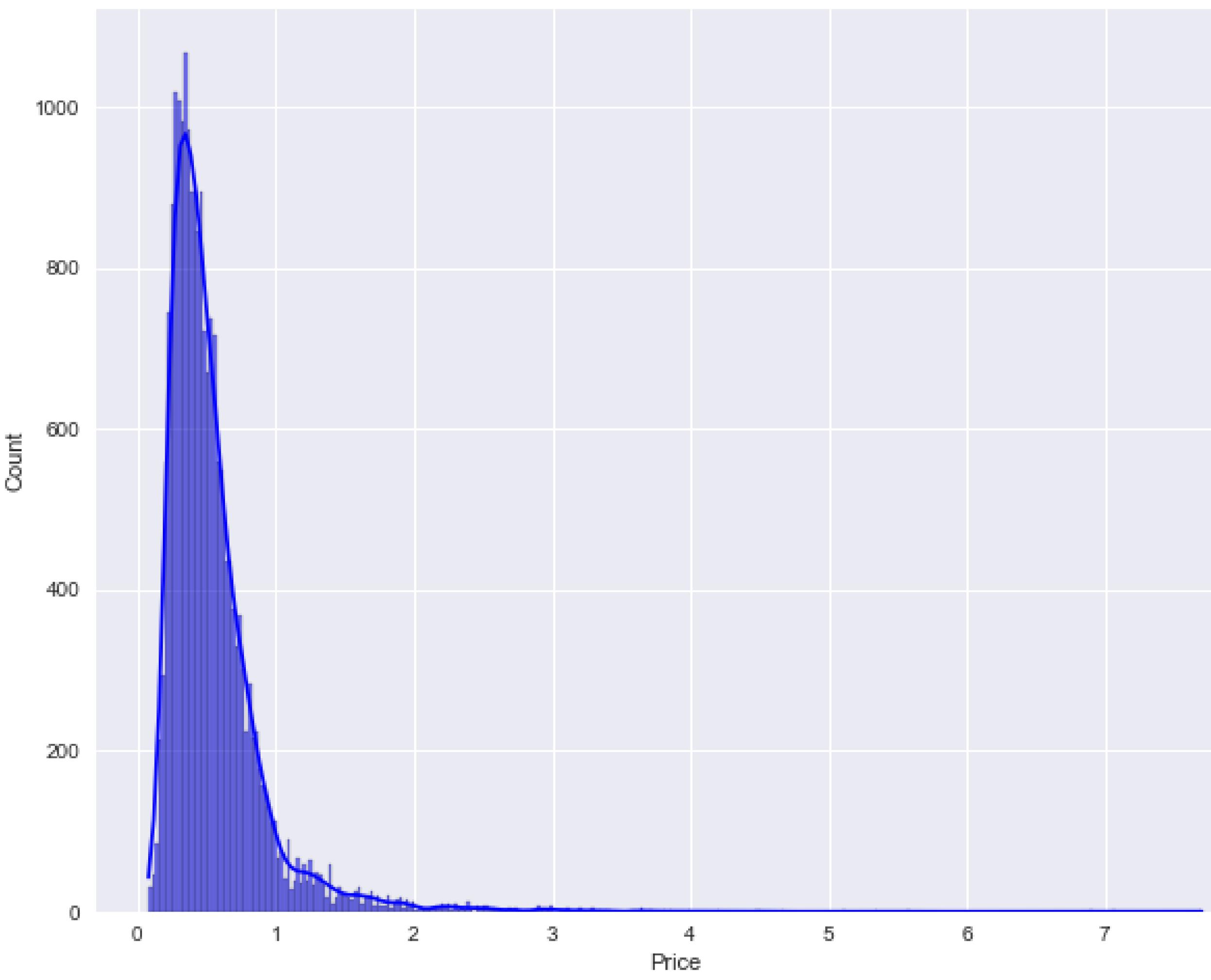
We use Scikit Learn together with StatsModels for creating the machine learning model

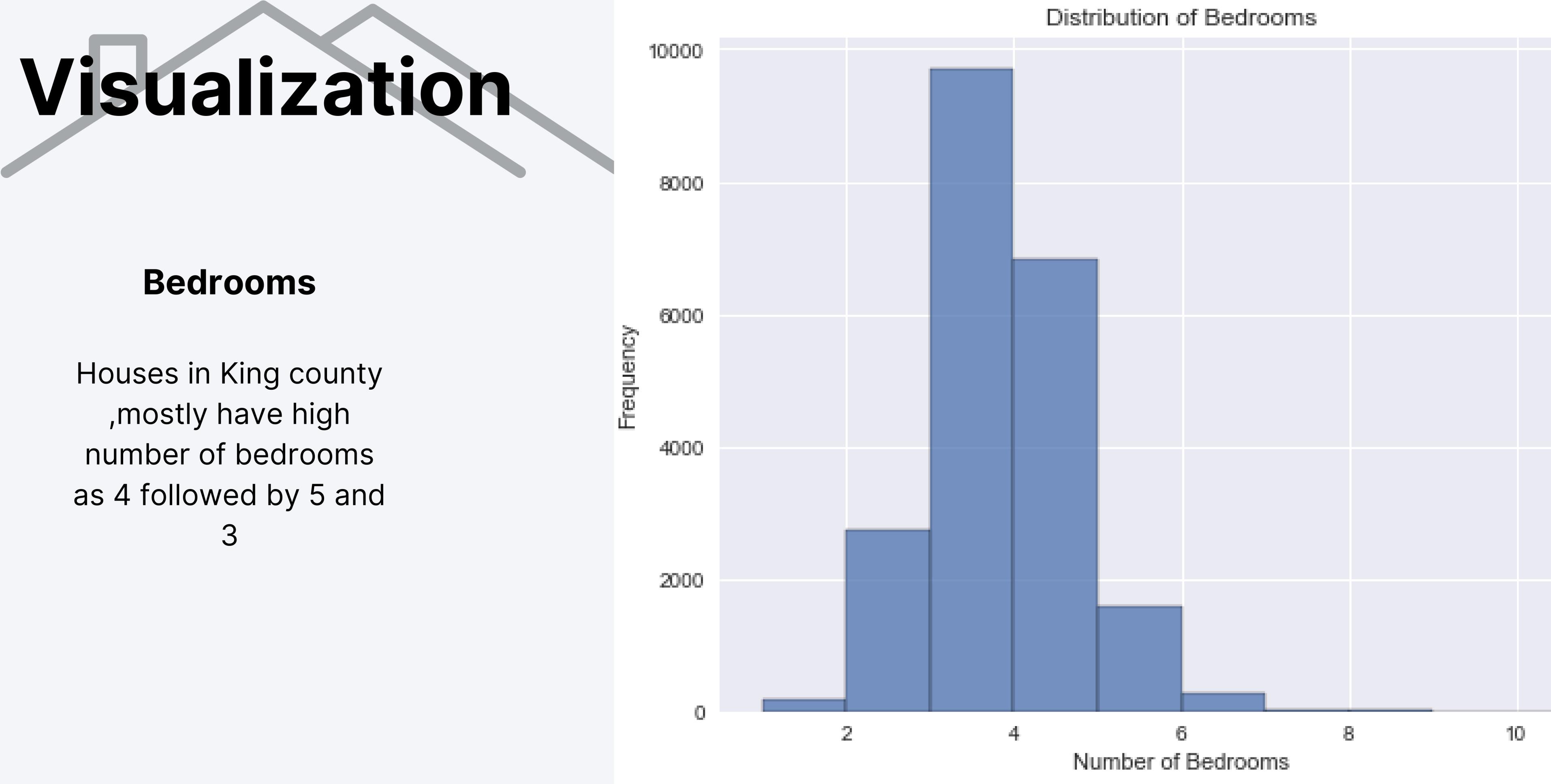


# Visualization

## The price distribution

There are houses that are overpriced  
in this region

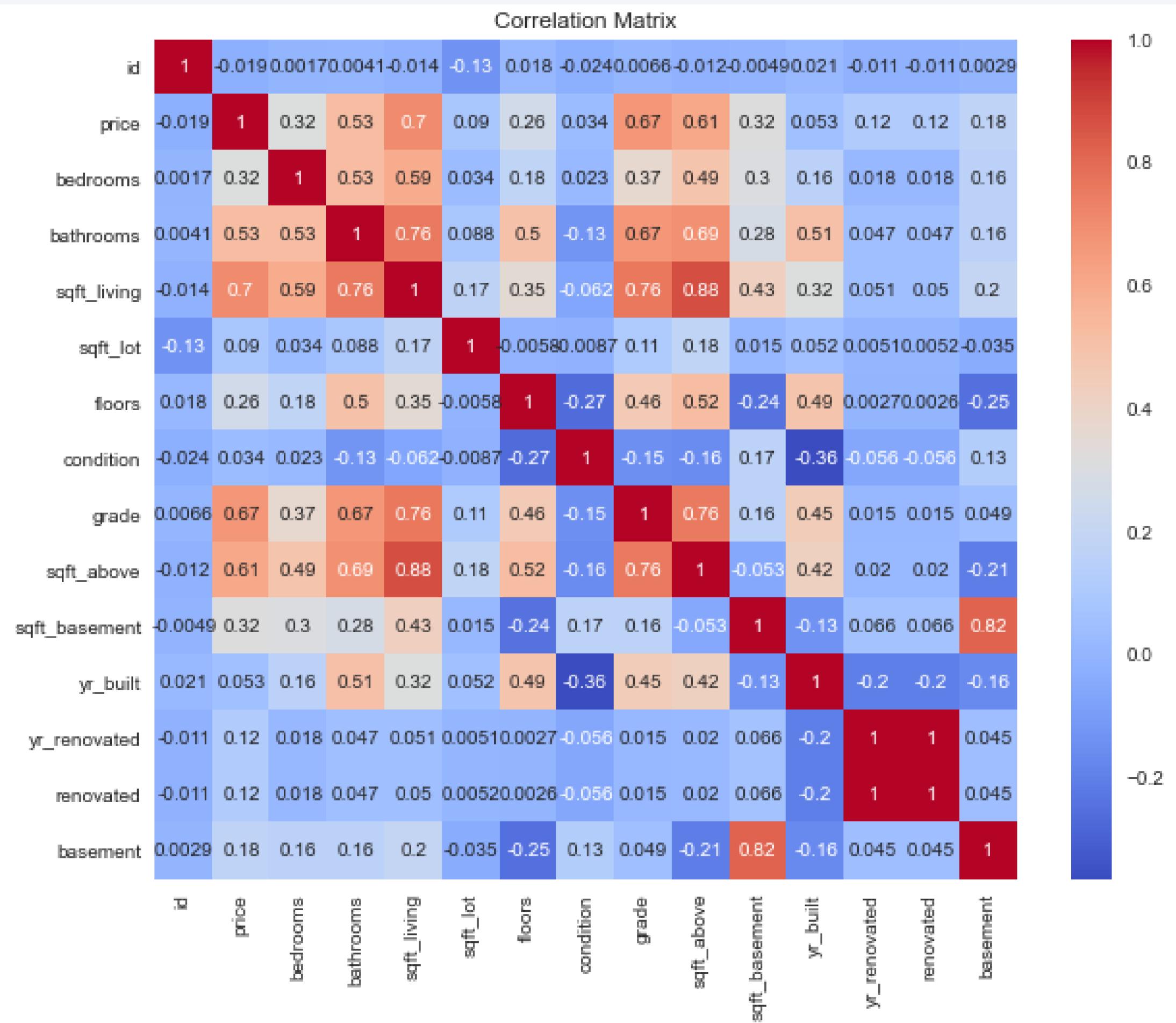




# Observations

There are features that have a higher correlation meaning that as they increase, the price also increases. Houses with more bathrooms, larger living spaces, higher grades, and more above-ground space tend to have higher prices.

The year built and square footage of lot do not really affect the price of a house because of weak correlation with the price.



# Modelling

## Baseline model

Adjusted R squared = 51%  
RMSE = \$264,562

The feature with the highest correlation with price is 'sqft\_living' and so we use that here.

## Multiple Regression Model

Adjusted R squared = 60%  
RMSE = \$239,783.  
Multiple features, including categorical variables like "view" and "waterfront."

## Polynomial Regression

Adjusted R squared = 64%  
RMSE = \$225,788

To address the skewed target variable, "price."

## Recursive Feature Elimination

Adjusted R squared = 61%  
RMSE = \$237,989

Used RFE to select the most impactful features, resulting in a model with seven selected features.



# Recommendations

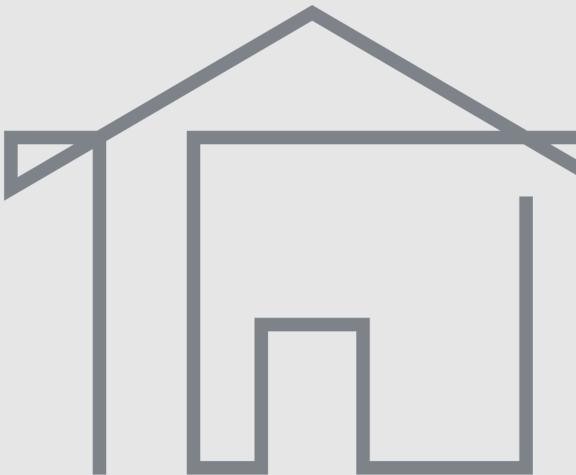


01

Invest in Key Features: Enhance property value by focusing on features with strong positive correlations, like adding more bathrooms, expanding living spaces, and improving the grade.

02

Condition is Less Influential: Don't overemphasize property condition in pricing decisions, as it shows no significant correlation with prices.



03

Strategic Renovations: Prioritize renovations that align with the preferences of the local market, rather than relying solely on year of construction or renovation.

04

Consult Experts: Collaborate with real estate professionals who understand the dynamics of King County's real estate market to make informed decisions.



# CONCLUSION



In conclusion, this project offers a comprehensive understanding of the complex King County real estate market in Washington.

Our objectives were to assess property values while providing actionable insights for homeowners regarding the impact of home renovations.



By employing data preparation, exploratory data analysis (EDA), and regression analysis, we addressed the challenges presented by a diverse and ever-changing market.

The results will not only benefit homeowners but also provide valuable guidance to real estate professionals in King County, helping them navigate this dynamic and diverse real estate market with confidence.



# Next steps

## 01 Employ more advanced techniques

To investigate variation of house price, it requires a more advanced modeling techniques such as regularization

## 02 Model Deployment

For practical use, whether for predictions, decision support, or other applications. Ensure that the model is integrated into your workflow or system as needed.



# Thank You



## QUESTIONS

>>>