

# KING COUNTY HOUSING PRICE PREDICTION MODEL

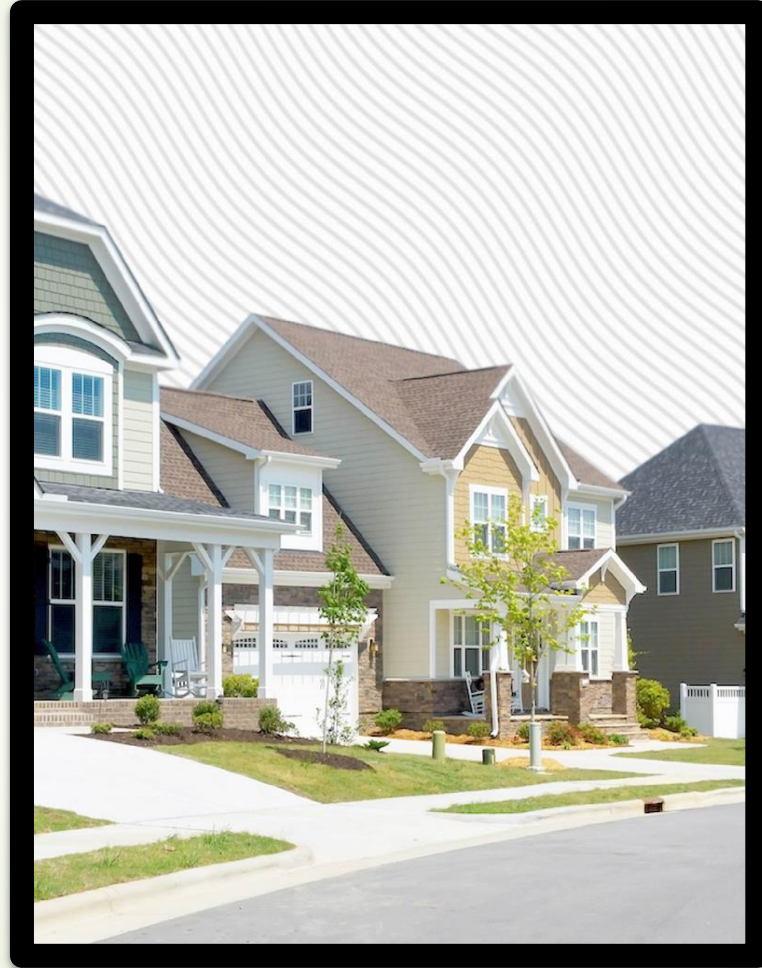
Data Driven solutions for the King County  
real estate market



# OVERVIEW

**This project was completed using data for the dataset provided by King County, Washington.**

**The goal for is to come up with a suitable model for a real estate company in Washington that will be used to predict house prices which will enable them maximize on the profits**





# Business Understanding

- **A company's real estate agent is curious about the elements that have a major influence on King County home values. This will help in formulating the optimal criterion to follow in order to optimize profit.**
- 

# PROBLEM STATEMENT

- **Given the complexity of factors that can affect the sale price of a house it is challenging to get a precise assessment**
- **Data driven analysis can provide key insight to agencies and clientele to help them make informed decisions**



# Objectives


- I. **Develop a predictive model for accurate house price estimations**
- II. **Identify key determinants that affect house pricing**
- III. **Investigate the impact of waterfront view on price**
- IV. **Evaluate the impact of House condition on price**
- V. **Determine the impact of house grade on price**





# METHODOLOGY AND TECHNIQUES

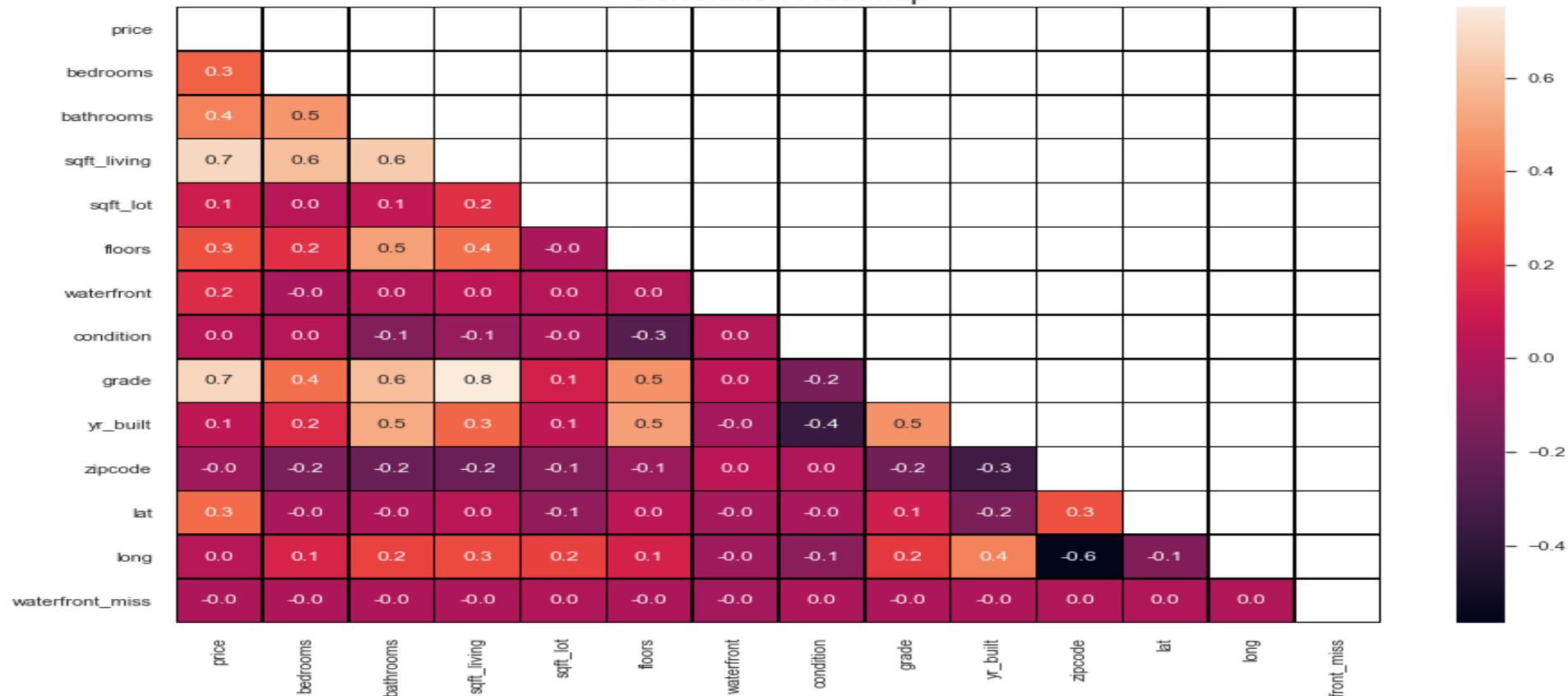
**The research used an iterative method to visualize attributes in connection to price, create a multiple linear regression model using Python and Stats Models to predict house sale prices, and discover characteristics that influence King County house prices.**



# DATA UNDERSTANDING

- There are 21597 records with 20 features in the King County data.
- Some null values in the waterfront, view, and year of renovation are present in the data.
- The highest-priced home sold for about \$7,700,000, while the average is almost \$540,000.
- The average house grade is 7, which indicates that the majority of sold homes have grades above average.
- In King County, the highest number of floors in a residence is 3.5.

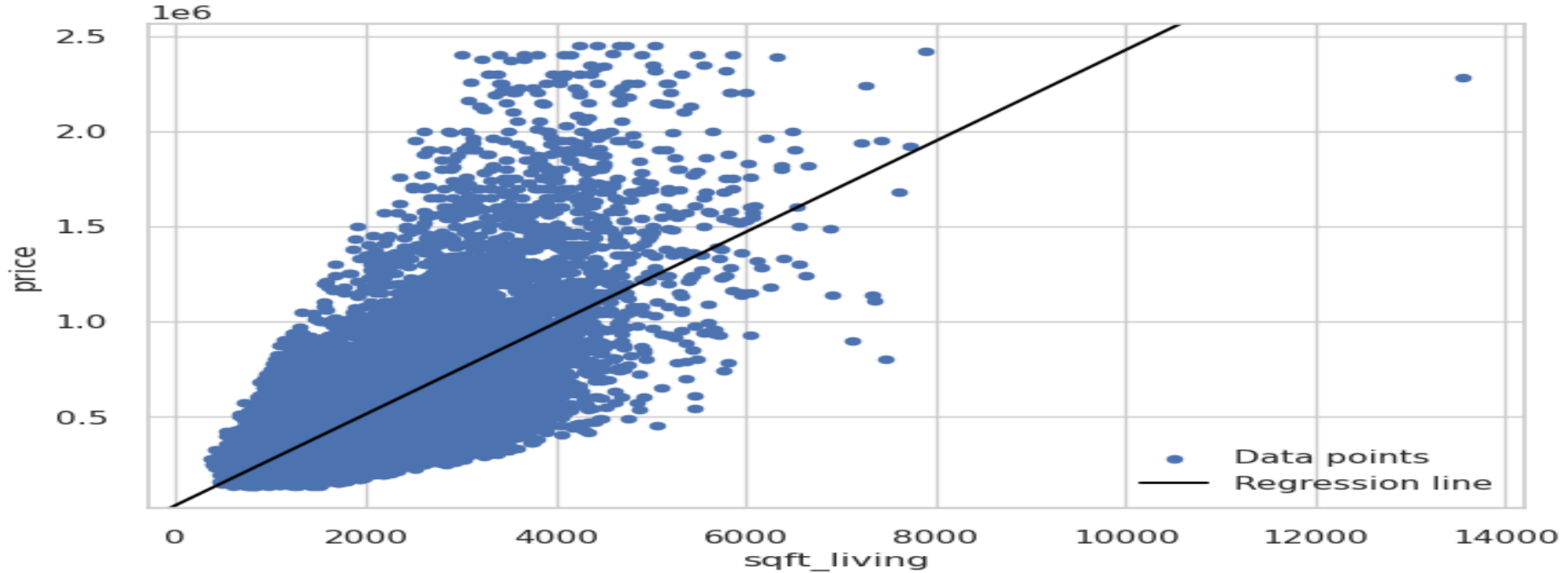
Correlation Heatmap



**Selected columns with significant correlations (bedrooms, bathrooms, sqft\_living, sqft\_lot, waterfront, grade, floors, condition) for modeling**



# Baseline Model



- linear regression model based on independent variable that has highest number of correlation with price
- Based on the model the living space increases price tends to increase by \$239 per square foot.

# CONCLUSION

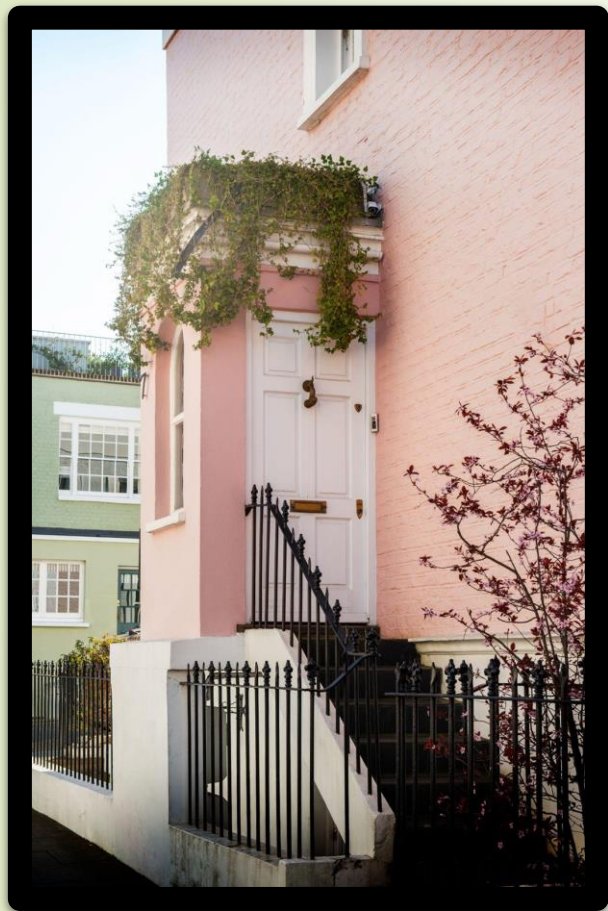
- The key factors are:
  - Waterfront location
  - Living area square footage
  - Overall grade of the property
- A house with the typical features for the area retails at \$528,969
- The data shows most houses of average condition which shows an investment potential for our stakeholders

# RECOMMENDATIONS

- **Different pricing strategies to emphasize Waterfront properties**
- **Investment in property upgrades**
- **Highlight spacious living areas**
- **Emphasis on Property Grade**

# NEXT STEPS

- **Comparison with other regression models**
- **Utilize supplementary data categories from the Kings County to improve the model's precision**
- **Use of current and long term data to understand trends in the market.**



Thank  
You