

# KINGS COUNTY HOUSING ANALYSIS

BY PRO  
REALTY





Branton Kieti



Priscilla Nzula



Beryl Saoke



Joseph  
Mangoka



Lorrah Ngine

## GROUP 10 TEAM MEMBERS

# OUTLINE



Summary



Business Problem



Stakeholder Objectives



Data



Modelling



Recommendations

# PROJECT OBJECTIVE

## **Project Introduction**

- Welcome to the Pro Realty Real Estate Investor House Price Prediction project.
- Objective: Develop a robust multiple linear regression model.

## **Predicting House Prices**

- Utilize machine learning techniques.
- Provide valuable insights into key factors impacting real estate prices.

# BUSINESS PROBLEM

## Pro Realty's Expansion Goals

Leading real estate firm with aspirations for growth.  
Focus on solidifying premier real estate investor status.



## Optimizing ROI

Data driven decisions for higher Return on Investment (ROI).  
Leveraging King County dataset.



## Strategic Decision-Making

Identifying lucrative investment opportunities.  
Maximizing ROI through informed expansion strategies.

# STAKEHOLDER OBJECTIVES



## House Pricing Factors

Identify critical variables influencing house prices.



## Predicting Pricing Trends

Forecast Kings County house prices with predictive analytics.



## Data-Driven Decisions

Empower buyers/investors with regression analysis insights.



# DATA

## King's County Housing Data

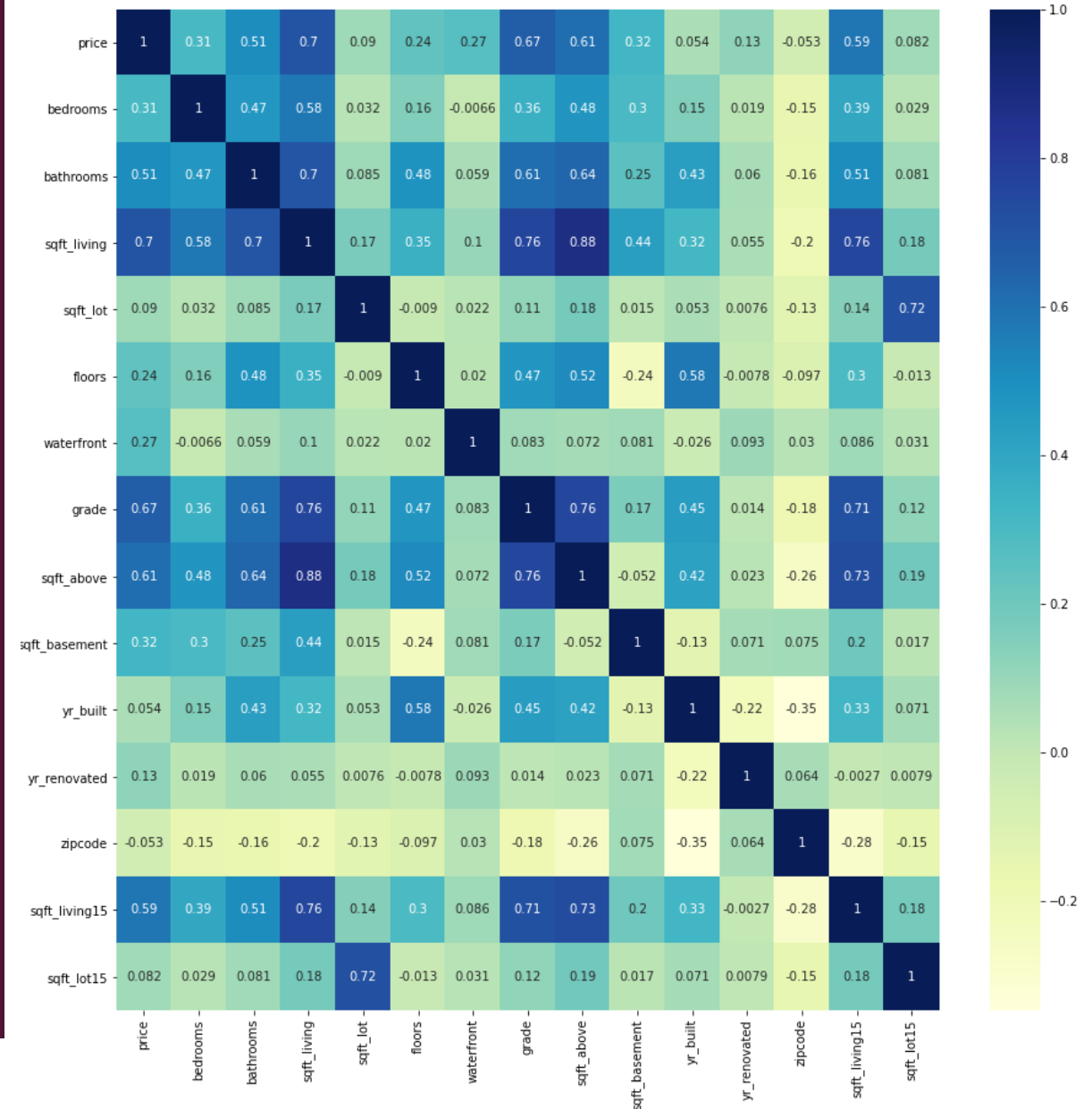
- Over 21,000 homes between 2014 and 2015
- Features 21 columns
- House Prices range from US\$78k to US\$7 Million





# DATA CORRELATION HEATMAP

- The correlation heatmap determines what features to use.
- Correlation Coefficients range from: -1 (strong negative correlation) to 1 (strong positive correlation)
- Strongly correlated Features:
  - Sqft\_living
  - Grade,
  - Sqft Above





# MODELLING

Category	Simple Linear Regression	Multiple Linear Regression
Mean Squared Error (MSE)	61,940,787,124.62	43,056,428,188.69
R-Squared ( $R^2$ )	47.92%	63.79%
Adjusted R-Squared	47.90%	63.71%
Mean Absolute Error (MAE)	170,780.93	137,762.19
Root Mean Squared Error (RMSE)	248,879.06	207,500.43

- Multiple linear regression outperforms the simple model:
  - Higher R-squared: Explains about 63.8% of house price variance.
  - Lower MSE and RMSE values: Improved accuracy.
  - Robust performance indicated by higher adjusted R-squared.

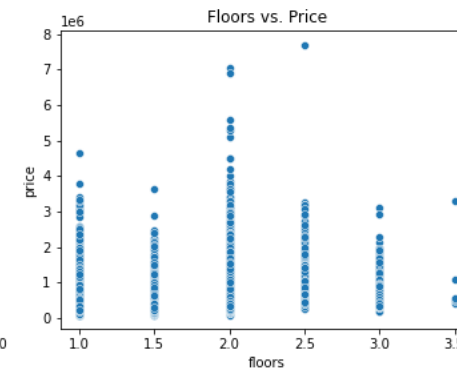
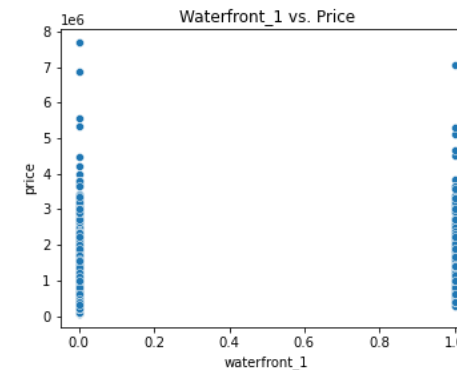
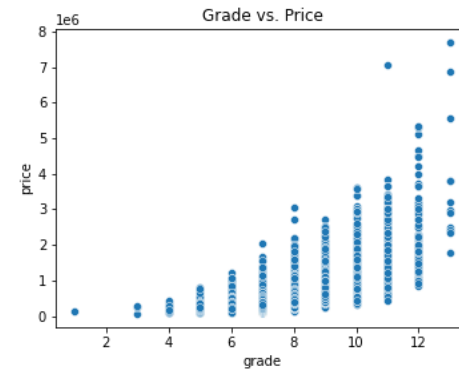
# RESIDUAL CALCULATIONS

## Y Intercept Value:

- Base value of a house with no additional features: \$6,594,806.17.

## Coefficients for House Features:

- Waterfront Property: +\$740,707.49
- Grade: +\$128,293.37 per grade unit.
- Bathrooms: +\$45,644.25 per bathroom.
- Living Area (sqft): +\$193.61 per sqft.
- Floors: +\$28,277.00 per floor.
- Bedrooms: -\$40,534.75 per bedroom



# RECOMMENDATIONS

## Waterfront Properties

- Significantly higher prices for waterfront properties.

## Grade

- Good construction quality is indicated by a higher grade.

## Bathrooms

- Properties with multiple bathrooms offer added convenience.

## Living Area (sqft)

- Spacious living areas meet diverse buyer needs and add value.

## Floors

- Additional floors are associated with premium pricing.

---

# THANK YOU

- Any Questions?

