

Predicting House Prices in King County : A Data-Driven Approach for Real Estate Optimization





Business Context & Problem Statement

How can King County real estate data be used to effectively predict house prices?

What strategic recommendations can help investors maximizing their profits and agencies to optimize pricing and investment decisions?

Dataset Overview & Preprocessing



Data Source :

King County Housing
Dataset (21,613 records)

Preprocessing Steps :

- Removed outliers via IQR on price
- Dropped irrelevant columns (e.g., id, date, zipcode, etc.)
- Excluded houses with >10 bedrooms

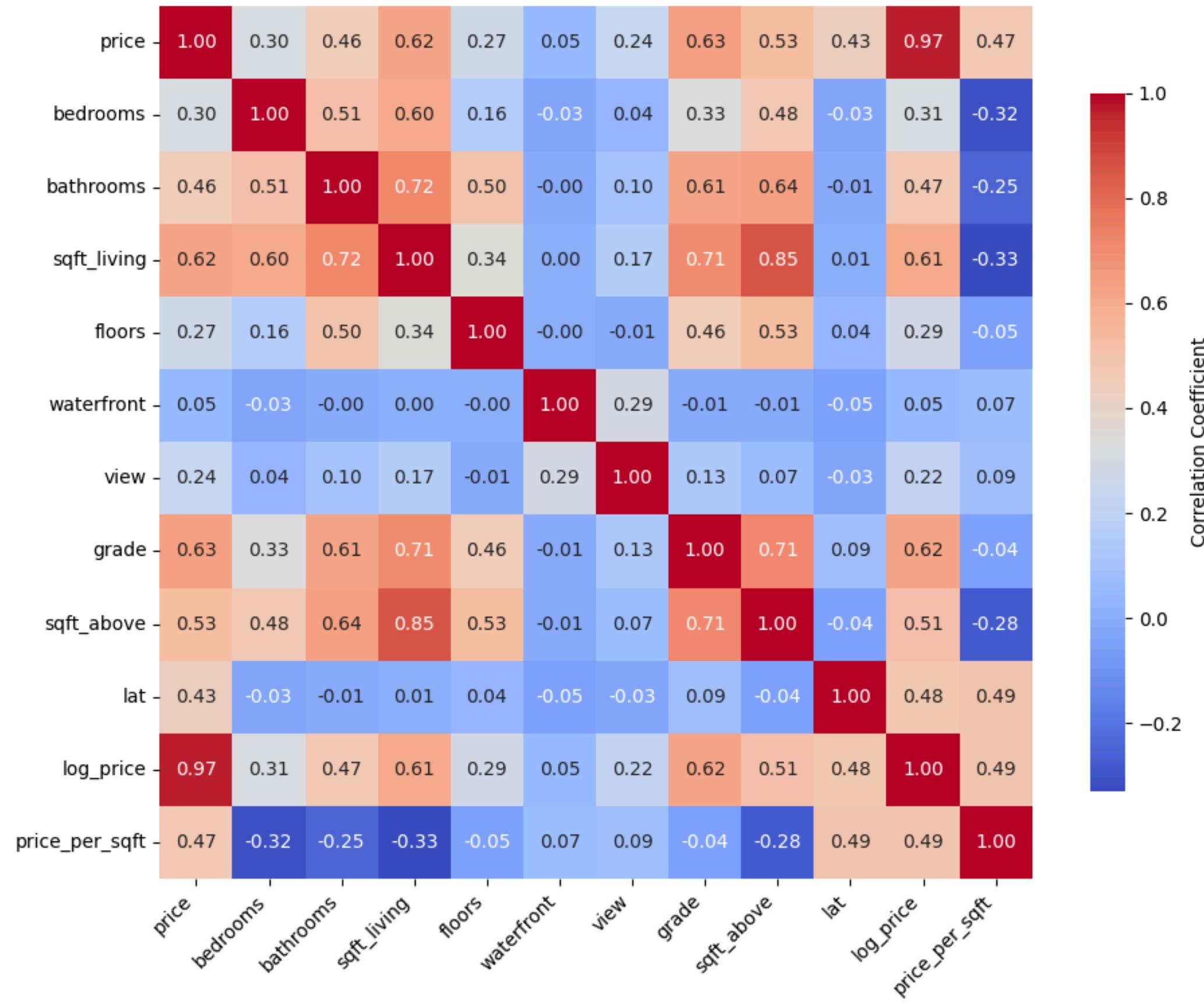
Methodology :

- Exploratory Data Analysis (EDA)
- Multiple Linear Regression for prediction

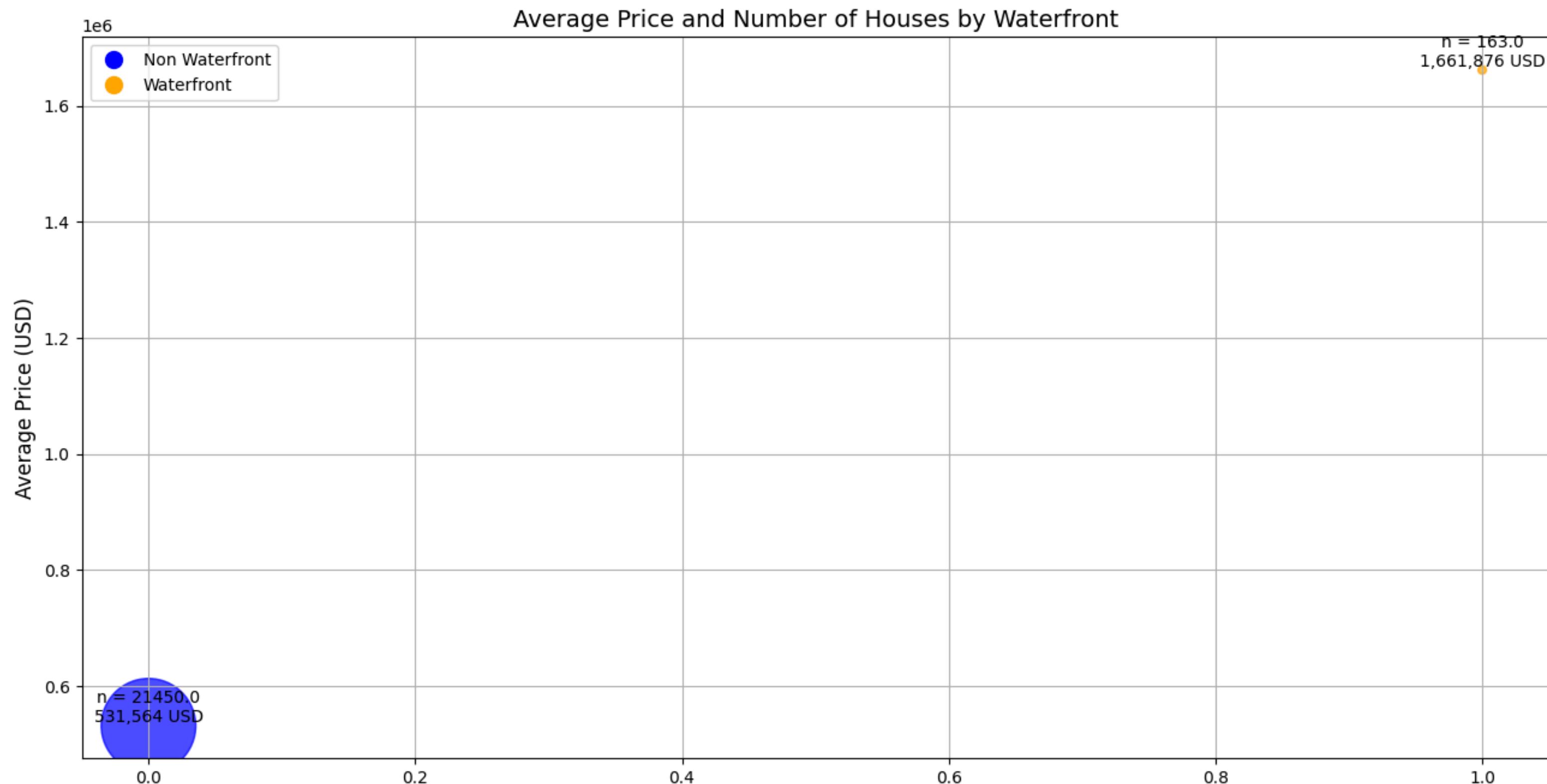
Exploratory Data Analysis (EDA)

- Strong correlations: sqft_living and grade with price
- No significant impact of waterfront properties on price

Correlation Matrix (Selected Columns)



EDA 2



Multiple linear regression & Prediction Results

Model Used :

Multiple Linear Regression
(predicting price)

Selected Features :

Grade, Sqft_living,
Bathrooms, Latitude,
Bedrooms, View,
Waterfront, Sqft_above

Performance Metrics :

- $R^2 \approx 0.66$ (log scale)
- RMSE (log_price) ≈ 0.26
- RMSE (Price in USD) $\approx \$137K$

Predictive Pricing Tool

- Inputs: Grade, Sqft_living, Bathrooms, Latitude, Bedrooms, View, Waterfront, Sqft_above
- Output: Predicted Sale Price (USD)

House Price Prediction

Grade :	<input type="text" value="7"/>
Sqft Living :	<input type="text" value="2000"/>
Bathrooms :	<input type="text" value="3"/>
Latitude :	<input type="text" value="47.5112"/>
Bedrooms :	<input type="text" value="3"/>
View (0 à 4) :	<input type="text" value="0"/>
Waterfront (0=No,1=Yes) :	<input type="text" value="0"/>
Sqft Above :	<input type="text" value="1800"/>

Predict

Price predicted : 379,711.88 USD

House Price Prediction

Grade :	<input type="text" value="8"/>
Sqft Living :	<input type="text" value="2000"/>
Bathrooms :	<input type="text" value="3"/>
Latitude :	<input type="text" value="47.5112"/>
Bedrooms :	<input type="text" value="3"/>
View (0 à 4) :	<input type="text" value="0"/>
Waterfront (0=No,1=Yes) :	<input type="text" value="1"/>
Sqft Above :	<input type="text" value="1800"/>

Predict

Price predicted : 601,257.13 USD

Strategic Recommendations

Improving Grade : A 1-point increase in grade leads to an estimated 14% increase in price.

Targeting Waterfront : Properties on the waterfront are valuated an additional (~+38%).

Location Factors : Higher latitude areas are associated with higher prices.

Recommendations :

- **For Investors :** Focus on properties with potential for grade improvement and undervalued properties
- **For Agencies :** Use the predictive tool to set competitive listing prices and identify undervalued properties.

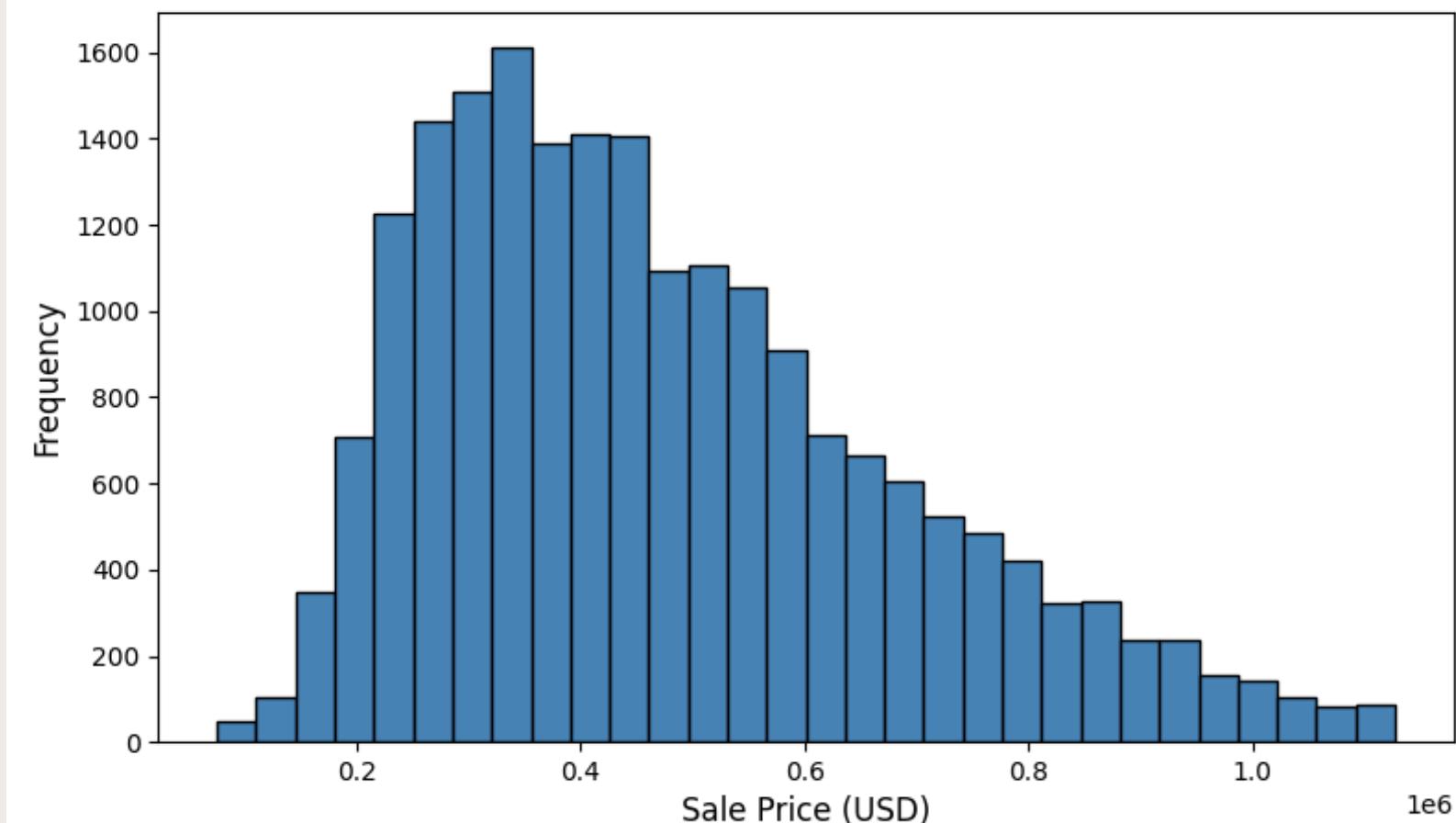
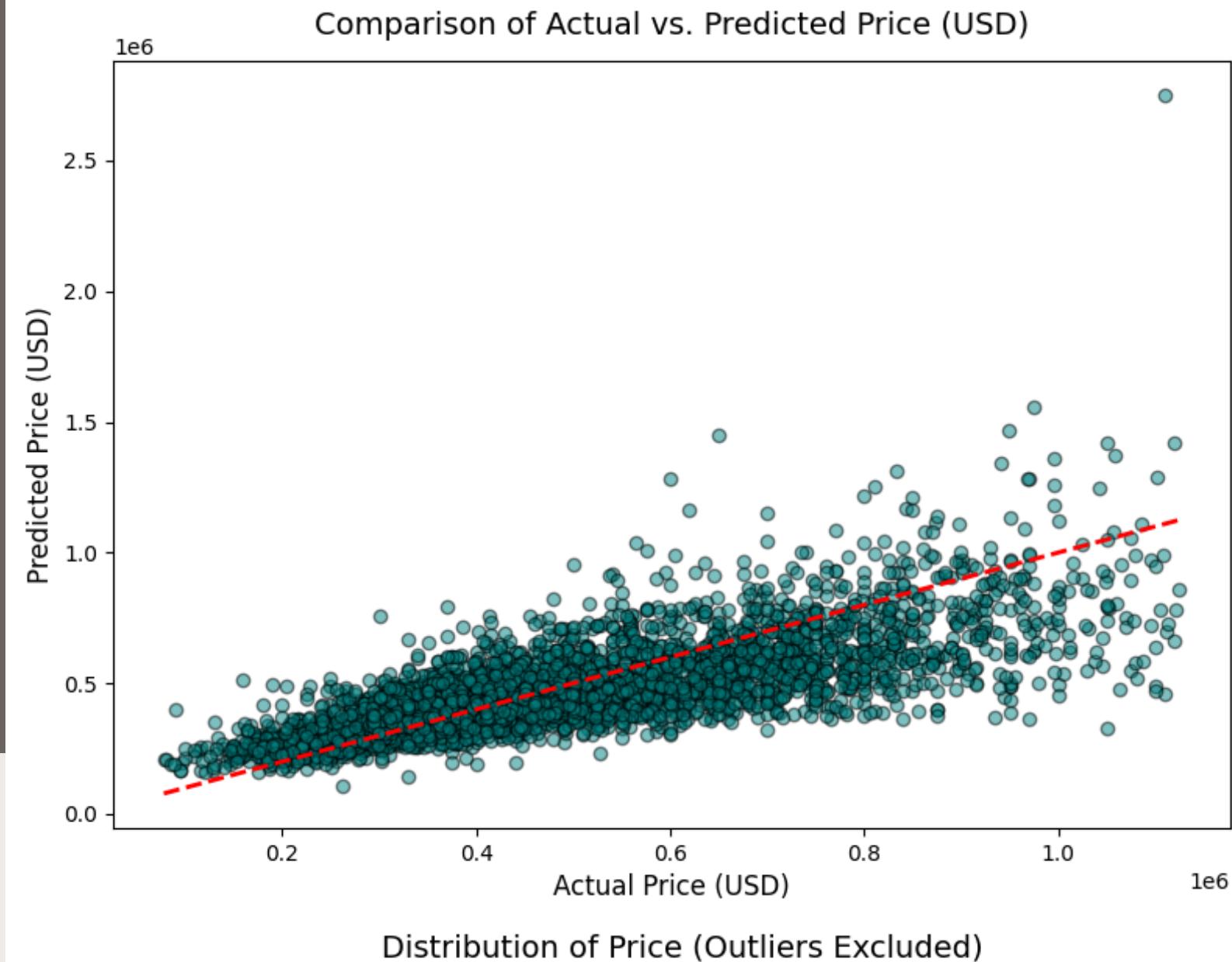
References and appendix :

Data sources :

<https://www.kaggle.com/datasets/harlfoxem/housesalesprediction>

<https://github.com/Shivanikanodia/House-Sales-in-King-Country>

- Tools & Libraries :
- Python (pandas, NumPy, scikit-learn, matplotlib, seaborn)



Thank You.

