SATELLITE IMAGERY–BASED BASED PROPERTY VALUATION VALUATION

A multimodal machine learning approach to predict residential property prices by combining structured housing data with satellite imagery.

Made By: Vansh Dhiman
Enrollment: 23115154
Domain: Data Science
Year: 2025–26

CHAPTER 1

THE PROBLEM WITH TRADITIONAL TRADITIONAL VALUATION

WHAT'S MISSING?

Traditional property valuation relies on structured data: square footage, bedrooms, bathrooms, and location coordinates. But these features miss critical context.

Two similar homes can differ dramatically in value based on neighborhood factors that tabular data can't capture.

THE GAP

- Proximity to water bodies
- Green space availability
- Road connectivity
- Urban density patterns
- Nearby infrastructure

These environmental factors strongly influence buyer perception and pricing.

WHY SATELLITE IMAGERY MATTERS

VISUAL CONTEXT

High-resolution satellite images encode patterns related to land use, vegetation, water proximity, and urban layout.

HUMAN INTUITION

Visual cues align with human perception of desirability and "curb appeal" in ways structured data cannot.

COMPLEMENTARY DATA

Imagery provides neighborhood-level information that enhances traditional property features.

CHAPTER 2

PROJECT OBJECTIVES

01

BASELINE MODEL

Build robust property valuation using structured tabular housing data.

02

IMAGE ACQUISITION

Programmatically acquire satellite images using geographic coordinates.

03

FEATURE EXTRACTION

Extract high-level visual features using pretrained CNN architecture.

04

MODEL DEVELOPMENT

Develop image-only and multimodal regression models for price prediction.

05

PERFORMANCE ANALYSIS

Compare tabular-only, image-only, and multimodal approaches.

06

EXPLAINABILITY

Apply Grad-CAM to understand which visual regions influence predictions.

Problem Statement & Objectives

Problem Statement :

Traditional housing data lacks the **environmental context** (neighborhood quality, green space, etc.) critical for accurate valuation.

This project aims to improve price predictions by developing a **multimodal pipeline** that fuses **tabular property data** with **satellite imagery**. The core challenge is effectively extracting visual features and integrating these diverse data types into a single, high-performance model.

Project Objectives :

The objectives of this project are as follows:

- To build a robust baseline property valuation model using only structured tabular housing data.
- To programmatically acquire satellite images using geographic coordinates (latitude and longitude) associated with each property.
- To extract high-level visual features from satellite images using a pretrained Convolutional Neural Network (CNN).

- To develop and evaluate image-only and multimodal regression models for property price prediction.
- To compare the performance of tabular-only, image-only, and multimodal models using standard regression metrics.
- To apply explainability techniques (Grad-CAM) to understand which visual regions in satellite images influence model predictions.

Scope and Expected Outcomes :

The scope of this project is limited to residential property valuation using historical housing transaction data and publicly available satellite imagery. The expected outcomes include a deeper understanding of the strengths and limitations of multimodal learning in real estate analytics, insights into how environmental context affects property prices, and a reproducible end-to-end pipeline that demonstrates the practical challenges of integrating visual data into traditional machine learning workflows.

Dataset Description

Data Source :

The primary dataset used in this project is a publicly available residential housing dataset containing historical property sale records. The dataset includes structured information describing property characteristics along with geographic coordinates (latitude and longitude) for each property. These coordinates enable the programmatic retrieval of satellite imagery corresponding to each location.

Tabular Data :

Category	Example Features
Size	sqft_living, sqft_lot
Quality	condition, grade
Location	lat, long
Neighborhood	sqft_living15, sqft_lot15

Satellite Imagery Data :

Satellite images were programmatically fetched using geographic coordinates to capture neighborhood-level context. Due to API and computational constraints, images were collected for a stratified subset of properties covering the full price distribution.

Data Preparation :

Tabular features were cleaned, log-transformed (price), and standardized prior to modeling. Satellite images were resized and normalized, then converted into fixed-length feature embeddings using a pretrained convolutional neural network. These embeddings enable efficient integration of visual information into downstream regression models.

CHAPTER 2

PROJECT OBJECTIVES

01

BASELINE MODEL

Build robust property valuation using structured tabular housing data.

02

IMAGE ACQUISITION

Programmatically acquire satellite images using geographic coordinates.

03

FEATURE EXTRACTION

Extract high-level visual features using pretrained CNN architecture.

04

MODEL DEVELOPMENT

Develop image-only and multimodal regression models for price prediction.

05

PERFORMANCE ANALYSIS

Compare tabular-only, image-only, and multimodal approaches.

06

EXPLAINABILITY

Apply Grad-CAM to understand which visual regions influence predictions.

DATASET OVERVIEW

TABULAR FEATURES

Size: sqft_living, sqft_lot

Quality: condition, grade

Location: lat, long

Neighborhood: sqft_living15, sqft_lot15

Structured data describes internal property characteristics and basic location indicators.

SATELLITE IMAGERY

Images programmatically fetched using coordinates to capture neighborhood context. Stratified subset collected covering full price distribution.

CHAPTER 3

KEY INSIGHTS FROM EDA

PRICE DISTRIBUTION

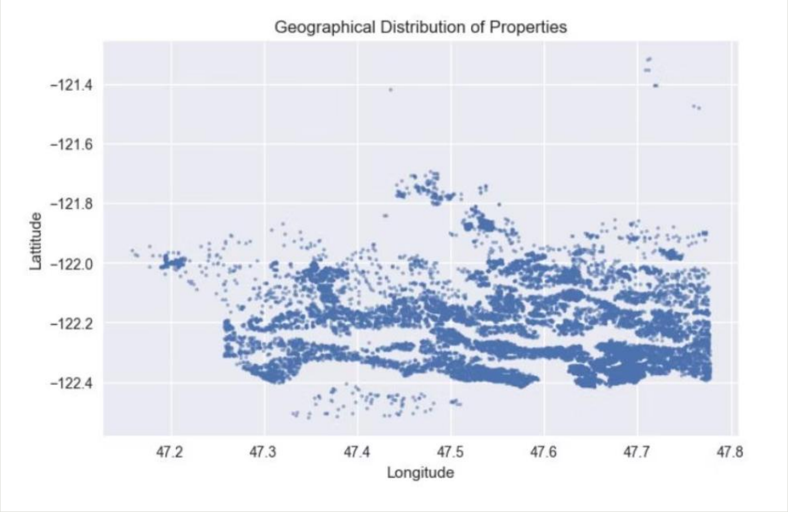
Highly right-skewed with few expensive properties. Log-transformation applied to stabilize variance.

FEATURE RELATIONSHIPS

Strong correlations: larger living areas, waterfront properties, and higher view ratings command premium prices.

GEOGRAPHIC CLUSTERING

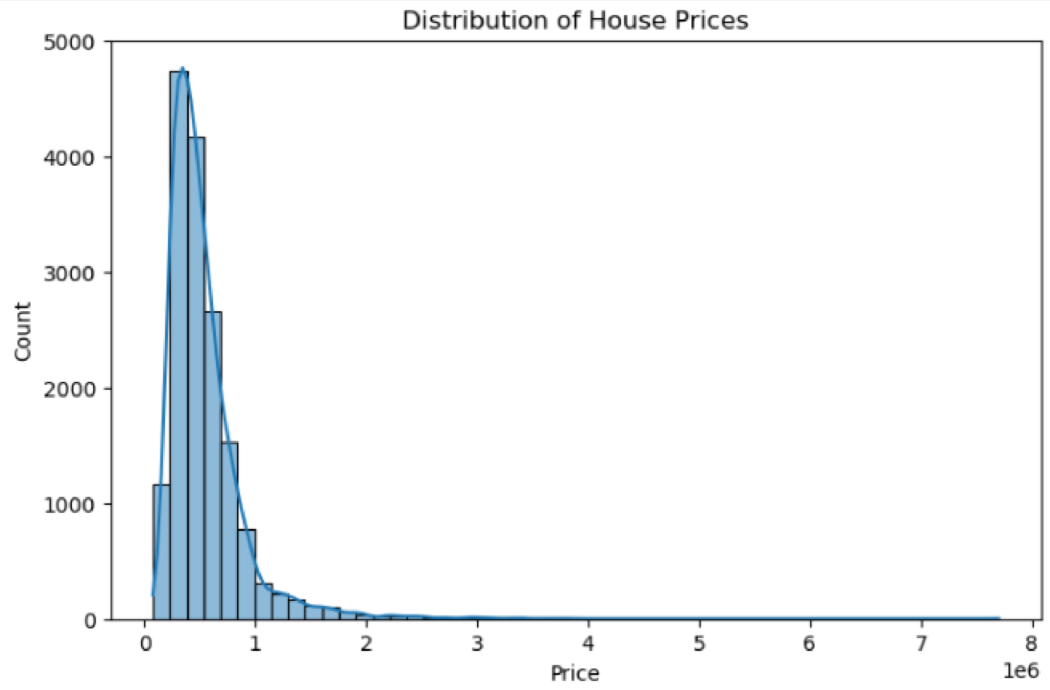
High-priced properties concentrate near water and premium zones. Lower-priced properties distributed inland.



KEY INSIGHTS FROM EDA

PRICE DISTRIBUTION

Highly right-skewed with few expensive properties.



Right-skewed distribution

Most houses are concentrated at lower to mid prices

A small number of very expensive houses stretch the tail to the right

This is called a positively skewed distribution.

Peak around ~300k–600k

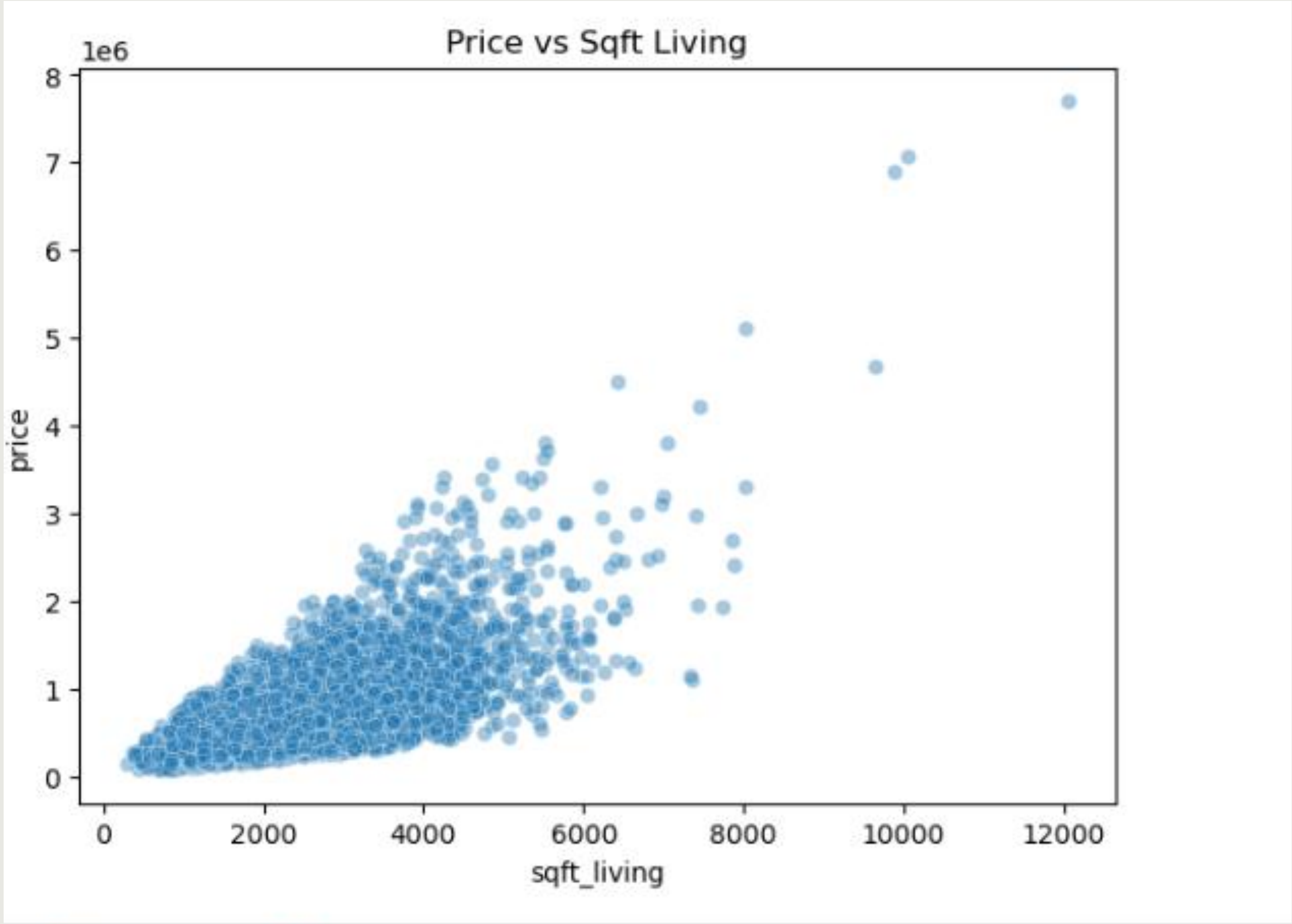
Long tail going up to ~7–8 million

KEY INSIGHTS FROM EDA

PRICE DISTRIBUTION

House price increases with **sqft_living**, showing a strong positive correlation..

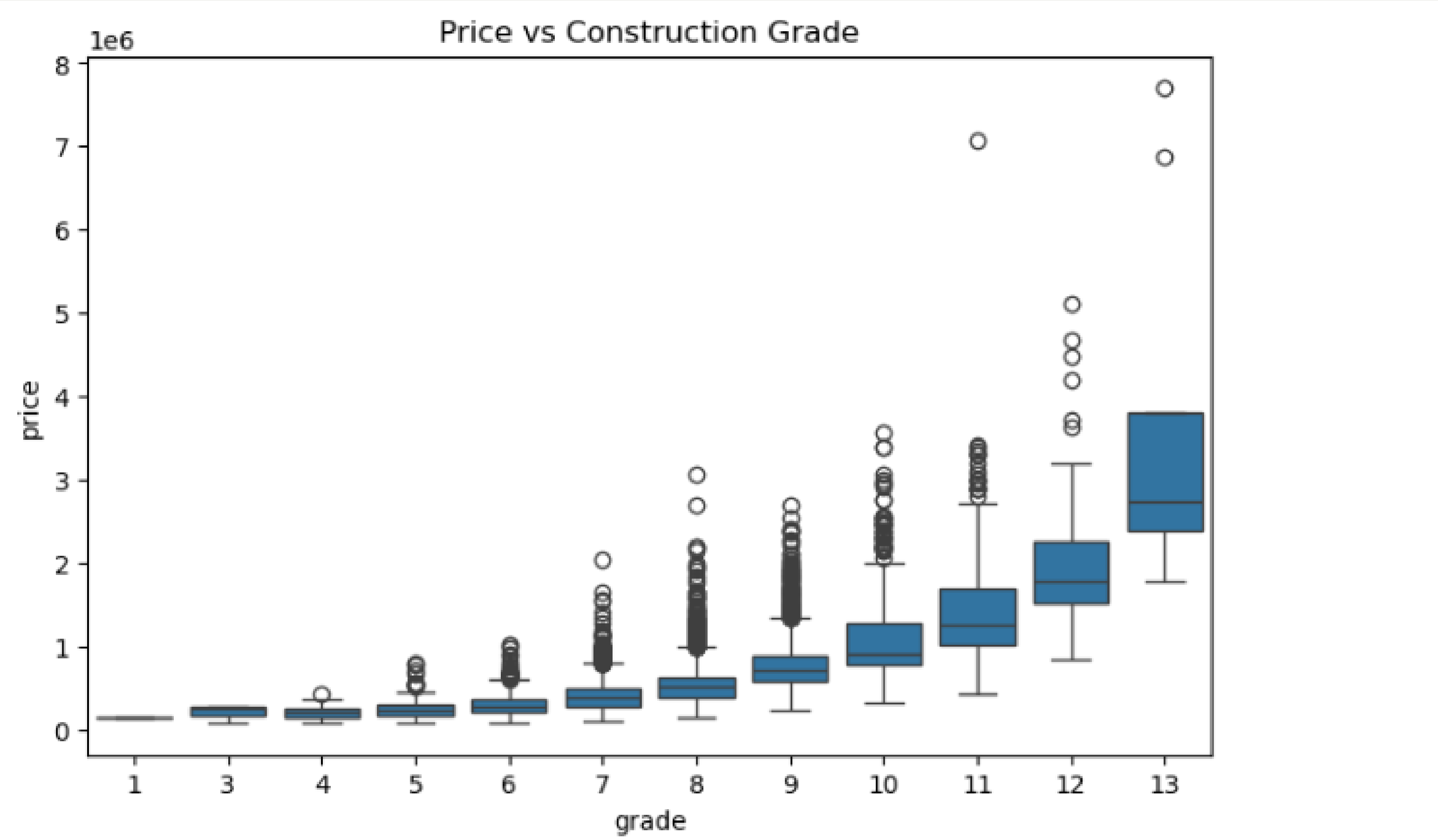
Larger homes generally have higher prices, but **price variability increases** at higher sizes.



KEY INSIGHTS FROM EDA

PRICE DISTRIBUTION

- House price increases **consistently with higher grade**.
- High-grade houses (11–13) show **significantly higher median prices**.
- Greater price spread at higher grades indicates **premium and luxury properties**.
- Grade is a **strong predictor** of house value.



KEY INSIGHTS FROM EDA

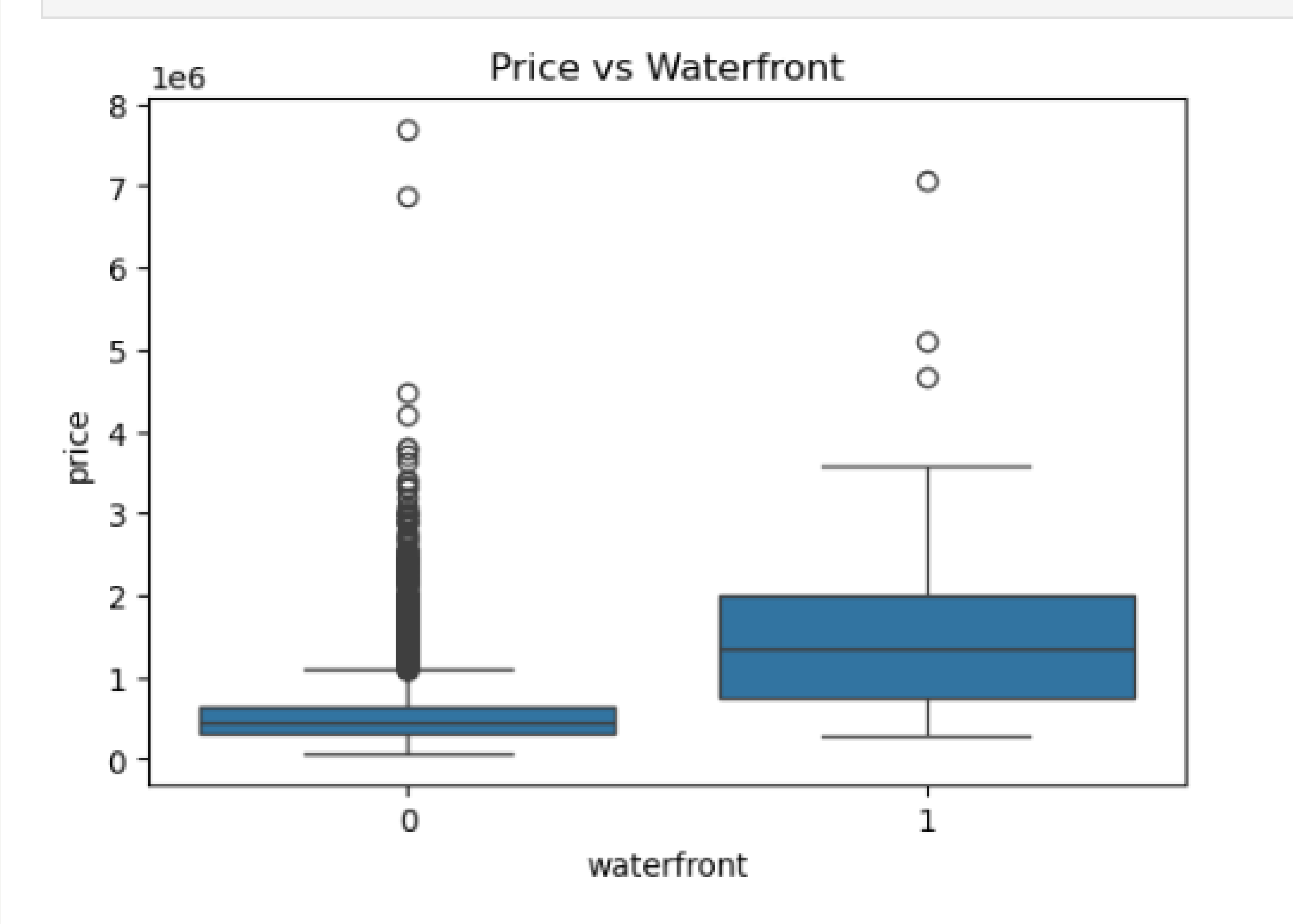
PRICE DISTRIBUTION

Waterfront properties have **much higher median prices** than non-waterfront ones.

Waterfront homes show **greater price variability**, indicating premium demand.

Even with fewer samples, waterfront has a **strong price impact**.

Confirms location-driven price premiums.



KEY INSIGHTS FROM EDA

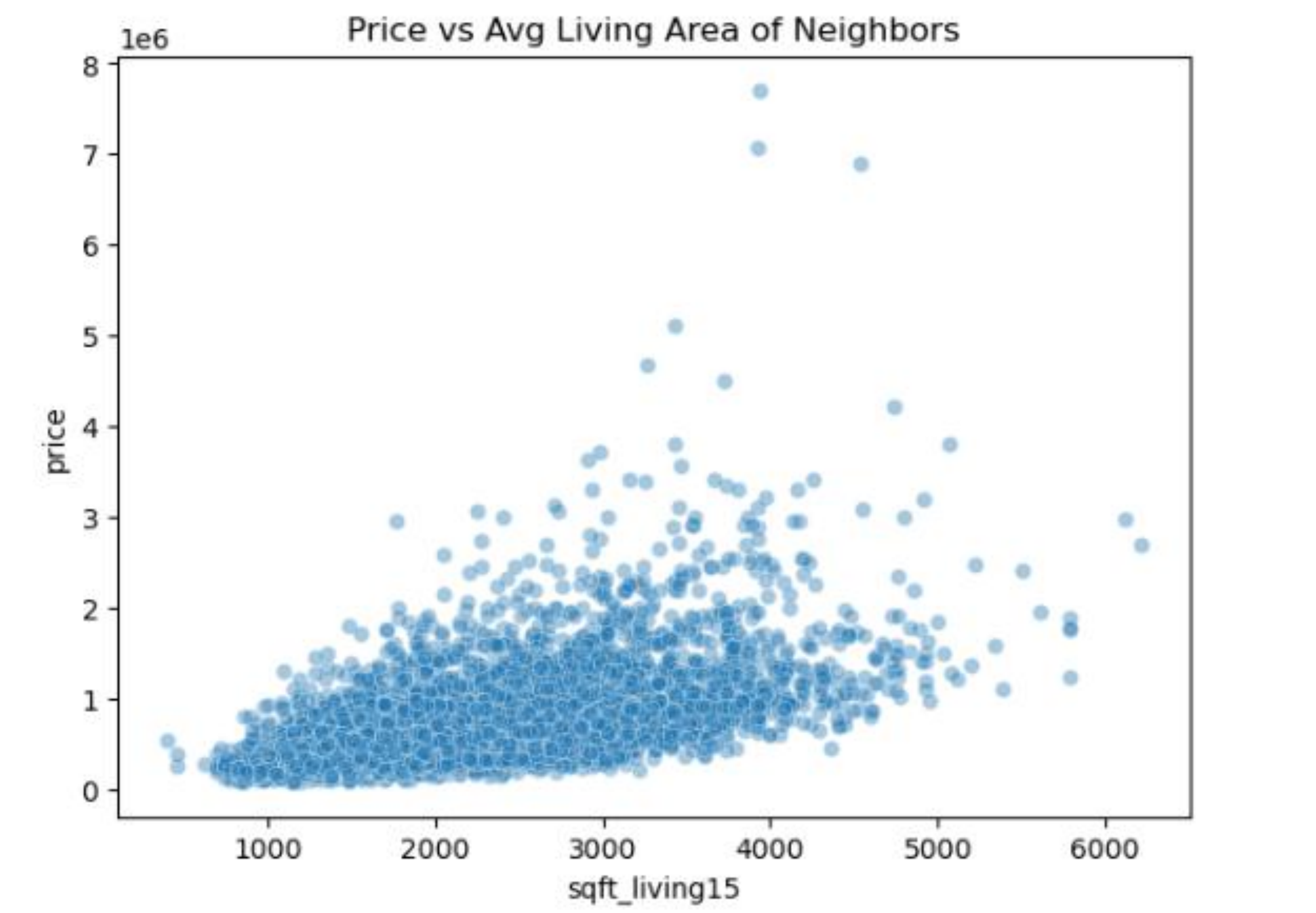
PRICE DISTRIBUTION

House prices increase with **larger neighboring house sizes**.

Indicates strong **neighborhood effect** on property valuation.

High variance suggests neighborhood alone does not fully determine price.

Supports using **environmental/contextual features**.

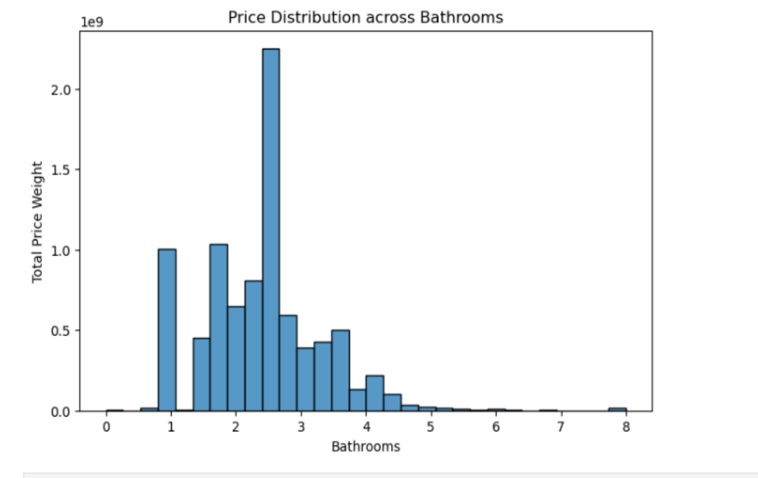


KEY INSIGHTS FROM EDA

PRICE DISTRIBUTION

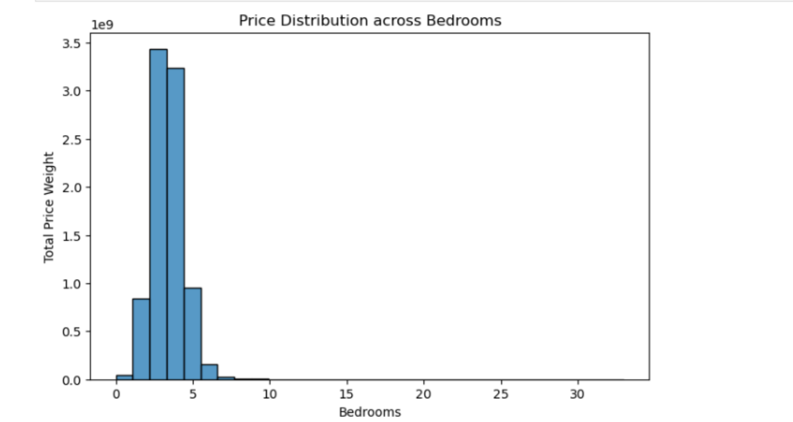
Price Distribution across Bathrooms

- Houses with **2–3 bathrooms contribute most to total property value**.
- Price generally increases with number of bathrooms.
- Very high bathroom counts are rare and act as **outliers**.
- Bathrooms are an important but **non-linear** feature.



Price Distribution across Bedrooms

- Most value contribution comes from **2–4 bedroom houses**.
- Increasing bedrooms beyond a point does **not guarantee higher prices**.
- Extremely high bedroom counts indicate **data anomalies or special properties**.
- Bedrooms alone are a **weak predictor** compared to size and grade.



Outlier Removal

Outlier Detection (Sqft per Bedroom)

- Some records show **unrealistic bedroom-to-area ratios**.
- Such anomalies can distort model training.
- Outliers were identified and filtered to **improve data quality**.

[273]:
```
df1[df1.sqft_living/df1.bedrooms<100]
```

[273]:

	price	bedrooms	bathrooms	sqft_living	floors	waterfront	view	co
3193	640000	33	1.75	1620	1.0	0	0	

[275]:
```
df2=df1[~(df1.sqft_living/df1.bedrooms<100)]
```

[277]:
```
df2
```

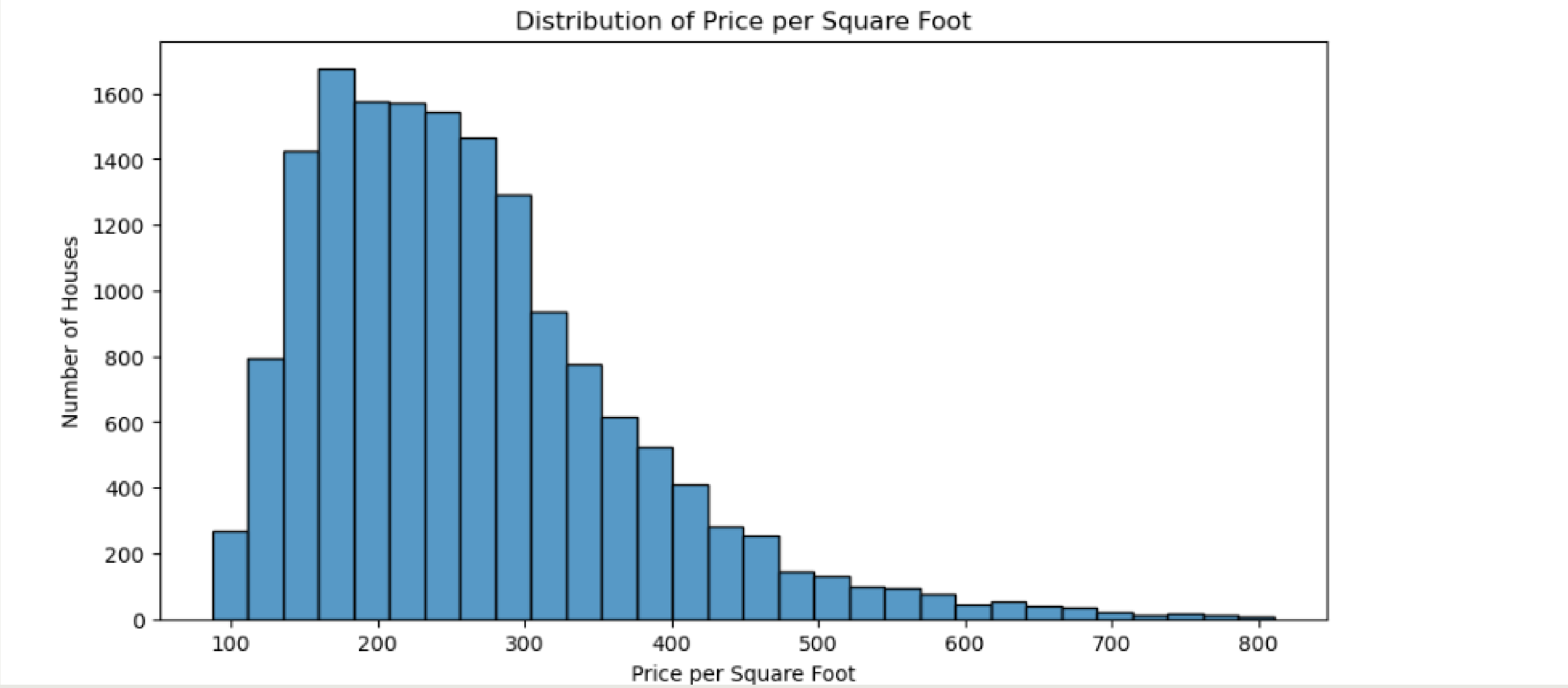
[277]:

	price	bedrooms	bathrooms	sqft_living	floors	waterfront	view	co
0	268643	4	2.25	1810	2.0	0	0	

Feature Engineering

EDA Insight – Price per Sqft (Living Area)

- Price per square foot varies widely across properties, indicating strong **location and neighborhood influence**.
- Properties with similar living areas can have very different price-per-sqft values, highlighting **quality and environmental factors**.
- Extremely high values correspond to **premium locations** (e.g., waterfront or high-grade areas).
- Normalizing price by living area reduces size bias and enables **fair comparison across properties**.
- This feature is effective for **outlier detection** and improves interpretability of pricing trends.



sqft_living15	price_per_square_feet
1660	148.421547
1720	153.125000
1870	116.279070
1240	284.273387
1590	181.250000
...	...
1000	378.000000
2927	129.559443
1690	271.226415

Final dataset

Final dataset exported

Price per Square Foot

```
[307]: df3.to_csv("final_cleaned_data.csv", index=False)

[ ]:
```

Modelling Approach

Tabular Baseline :

A baseline model was first built using structured tabular features to establish reference performance. The target variable (price) was log-transformed to reduce skewness, and numerical features were standardized. A gradient boosting regressor was used due to its strong performance on structured data and ability to model non-linear relationships.

Image Feature Extraction :

Satellite images were converted into numerical representations using a pretrained ResNet-18 model. The final classification layer was removed to extract fixed-length (512-dimensional) image embeddings. All CNN weights were frozen during feature extraction to reduce overfitting and computational cost.

Multimodal Fusion :

A simple early-fusion strategy was adopted, where standardized tabular features and image embeddings were concatenated and passed through a fully connected neural network for price prediction. This approach was chosen for its simplicity and interpretability.

Tabular Features ──▶ Tabular Encoder ─┐

├─▶ Fusion MLP ─▶ Price

Satellite Image ─▶ CNN ─▶ Embedding ─┘

Training Setup :

Neural network models were trained using mean squared error loss on log-transformed prices and optimized with Adam. A validation split and early stopping were used to control overfitting.

MULTIMODAL ARCHITECTURE

MULTIMODAL PRICE PREDICTOR

PRICE PREDICTION

Log-target output

TABULAR FEATURES

Standardized numeric inputs

FUSION MLP

Concatenate and transform

TABULAR ENCODER

MLP embedding layer

IMAGE CNN

Frozen CNN extractor

SATELLITE IMAGES

ResNet-18 512-d embeddings

TABULAR BASELINE

Gradient boosting on standardized features with log-transformed price target.

IMAGE FEATURES

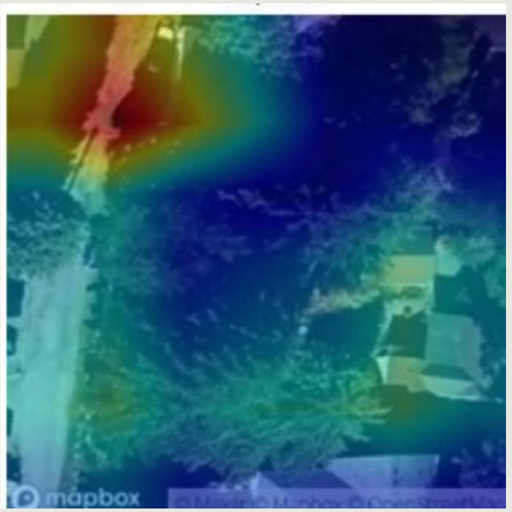
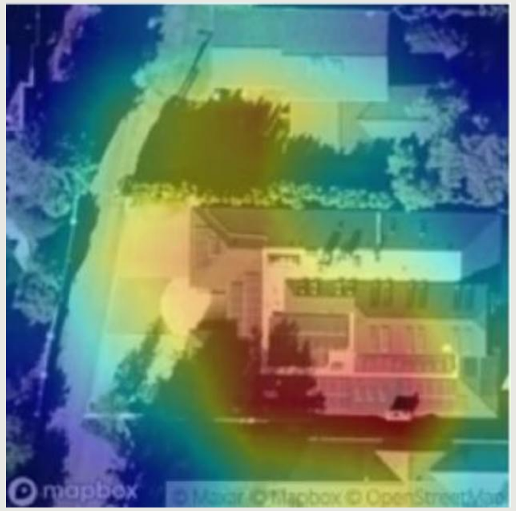
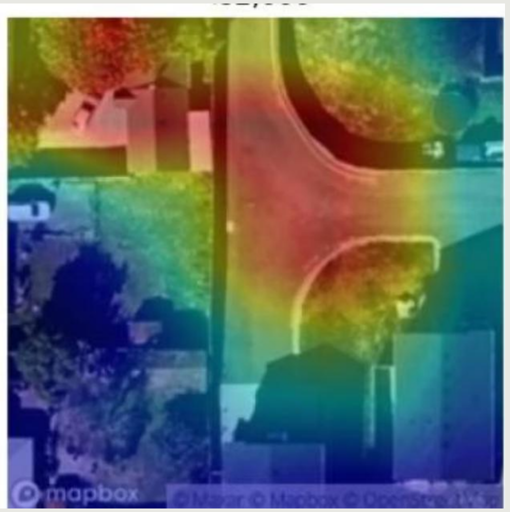
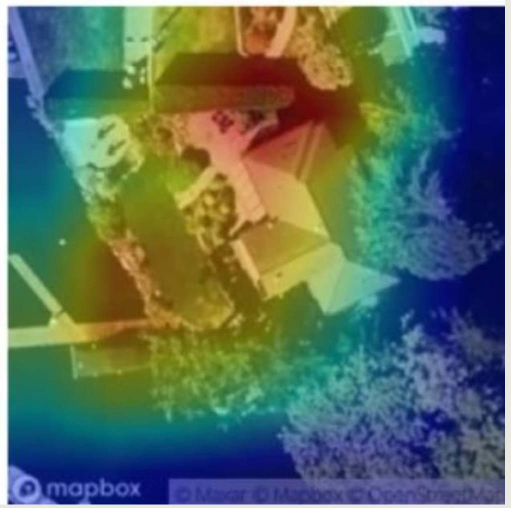
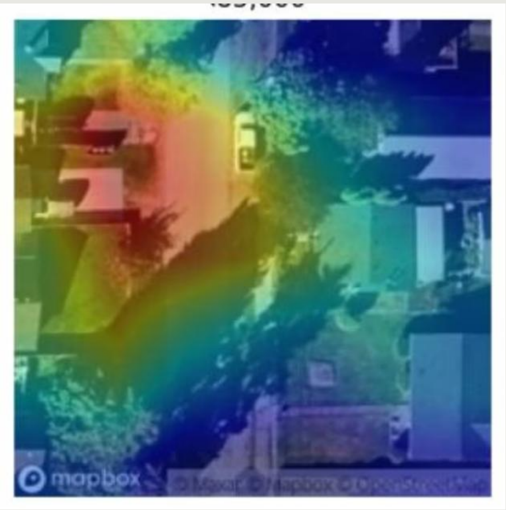
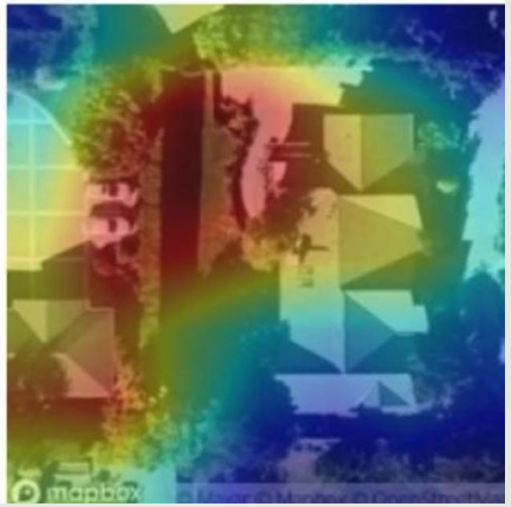
ResNet-18 extracts 512-dimensional embeddings with frozen weights.

EARLY FUSION

Concatenated features passed through fully connected network.

MODEL EXPLAINABILITY WITH GRAD-CAM

Gradient-weighted Class Activation Mapping reveals which image regions influence predictions, validating that the model focuses on meaningful neighborhood features.

High-value properties activate regions with water bodies, green spaces, and organized layouts. Low-value properties highlight dense urban structures with limited greenery.

Results & Model Comparison

Evaluation Metrics :

Model performance was evaluated using **Root Mean Squared Error (RMSE)** and **R^2 score** on a held-out validation set. RMSE measures prediction error magnitude in the original price scale, while R^2 indicates the proportion of variance explained by the model. For neural network models, predictions were made on log-transformed prices and converted back to the original scale for evaluation.

[1]:

```python
# Imports & metric helpers
import numpy as np
import pandas as pd
from sklearn.metrics import mean_squared_error, r2_score
def evaluate(y_true, y_pred):
    rmse = mean_squared_error(y_true, y_pred, squared=False)
    r2 = r2_score(y_true, y_pred)
    return rmse, r2



# Tabular-only baseline
tabular_rmse = 63380.2764
tabular_r2 = 0.9710



# Multimodal fusion (best attempt )
fusion_rmse = 803465.81566
fusion_r2 = -0.628686


results_df = pd.DataFrame({
    "Model": ["Tabular Only", "Multimodal Fusion"],
    "RMSE": [tabular_rmse, fusion_rmse],
    "R2": [tabular_r2, fusion_r2]
})

results_df
```

	Model	RMSE	R2
0	Tabular Only	63380.27640	0.971000
1	Multimodal Fusion	803465.81566	-0.628686

```python
# Tabular predicts better than the multimodal fusion
```

```python
# reason for low performance could be low dataset of images due to system limitation and higher correlation between different columns of tabular data anc
```

Discussion of Results :

The tabular-only model achieved the strongest performance, showing that structured property attributes capture most of the predictive signal. The image-only model performed worse than the tabular baseline but better than random, indicating that satellite imagery contains useful information. The multimodal fusion model did not outperform the tabular model, suggesting limitations of simple early-fusion strategies.

CONCLUSION & FUTURE DIRECTIONS

KEY TAKEAWAYS

This project demonstrates both potential and limitations of satellite imagery for property valuation.

- Structured features remain primary predictors
- Satellite imagery captures meaningful neighborhood context
- Simple fusion strategies need refinement
- Grad-CAM confirms interpretable visual patterns

FUTURE WORK

- Advanced fusion architectures (attention-based, late-fusion)
- Fine-tune CNN on satellite imagery
- Incorporate temporal data and seasonal effects
- Higher-resolution imagery
- Additional spatial features

GitHub:https://github.com/vansh1435/Satellite-Imagery-Property-Valuation/tree/main

Thank You