🏠 Satellite Imagery Based Property Valuation Project:

-Arun Kaarthikeyan R

23321006

Introduction:

Predicting the market value of residential properties is an important problem in real estate, banking, and investment analysis. Traditionally, house price prediction models rely on tabular data, such as the size of the house, number of rooms, construction quality, and location-related attributes. While these features are useful, they do not fully capture the visual and environmental aspects of a property's surroundings.

For example, two houses with comparable size and structure can have quite different values depending on factors like greenery, road connectivity, or proximity to water bodies. These factors are often visible from satellite imagery but are difficult to represent using numbers alone.

In this project, a multimodal regression framework is developed to predict house prices by combining tabular housing data with satellite images. The goal is to study whether adding visual neighbourhood context can improve property valuation compared to a tabular-only model.

Dataset Description:

The tabular dataset used in this project is taken from the **House Sales Prediction Dataset**, which contains historical data for residential properties. Each record represents one house along with its attributes and sale price.

Target Variable

- price: The market value of the property.

Key Features

- Structural features such as bedrooms, bathrooms, and sqft_living.
- Land-related features like sqft_lot.
- Neighbourhood-level features including sqft_living15 and sqft_lot15, which represent the average size of nearby houses.
- Quality-related features such as condition and grade
- Location-based attributes like view, waterfront, latitude, and longitude

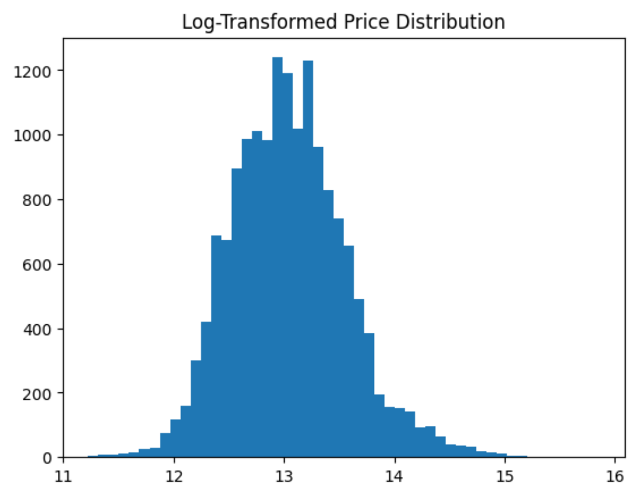
The latitude and longitude values are especially important, as they allow satellite images to be fetched for each property.

Exploratory Data Analysis (EDA):

Exploratory Data Analysis was performed to better understand the dataset and identify important patterns.

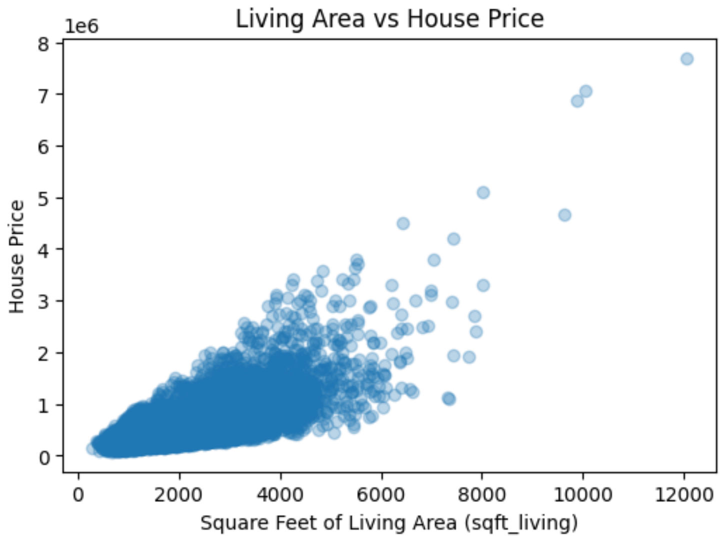
1.Price Distribution:

The distribution of house prices is highly skewed, with a few properties priced significantly higher than the rest. To address this, the target variable was **log-transformed**, which results in a more balanced distribution and improves model performance.



2.Living Area vs Price:

A clear positive relationship is observed between sqft_living and house price. As expected, houses with larger living areas have higher prices.

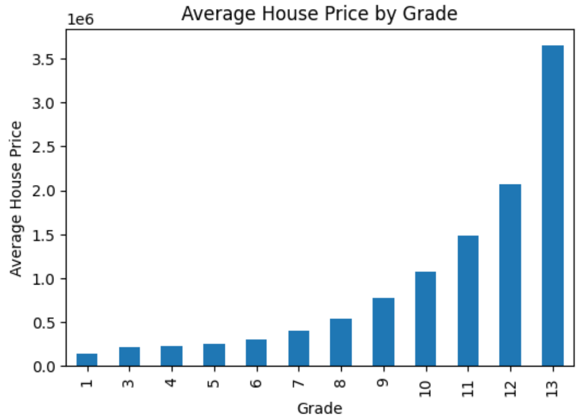


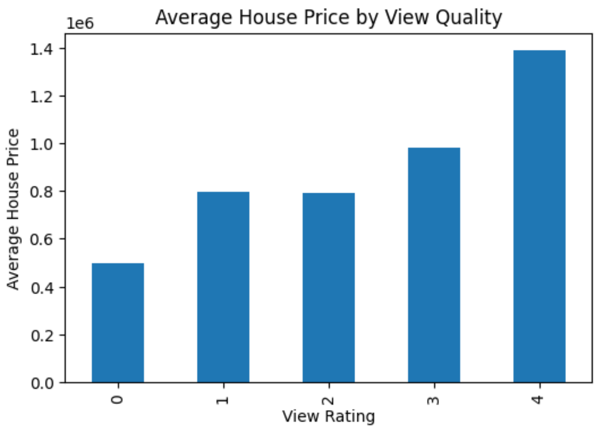
3.Neighborhood Effects:

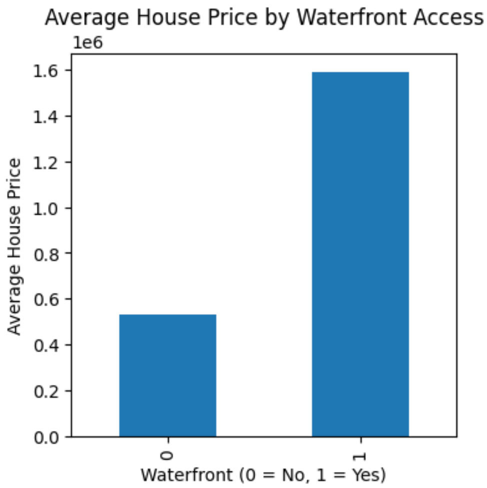
Neighbourhood-level features such as sqft_living15 and sqft_lot15 show that houses located in areas with larger neighbouring homes and lots tend to be more expensive. This highlights the importance of **local context** in determining property value.

4.Categorical Features:

- Higher construction grades correspond to higher prices.
- Houses with better views or waterfront access show a clear price premium.







Satellite Imagery and Visual Analysis:

To capture environmental and neighbourhood context, satellite images were fetched using **latitude and longitude coordinates** through the **Mapbox Static Images API**. Each image represents the area surrounding a property and provides visual information that is not available in tabular form.

These images capture features such as:

- Green spaces and tree cover
- Road networks and urban density
- Nearby water bodies
- Overall neighbourhood layout

Due to API and computational limitations, a **representative subset** of satellite images was used for multimodal modelling.

Methodology

1.Data Preprocessing

The tabular data was cleaned by handling missing values using median imputation. The target variable was log-transformed, and the dataset was split into training and validation sets. Numerical features were standardized to ensure stable model training.

2.Tabular Baseline Model

A Linear Regression model was trained using only the tabular features. This model serves as a baseline to evaluate how much improvement can be achieved by adding satellite imagery.

3.CNN-Based Visual Feature Extraction

Satellite images were passed through a pretrained ResNet-18 convolutional neural network. The final classification layer was removed, and the activations from the penultimate layer were used as 512-dimensional image embeddings. These embeddings capture high-level visual patterns in the neighbourhood environment.

This transfer learning approach avoids the need to train a CNN from scratch and makes efficient use of pretrained knowledge.

4.Multimodal Fusion

The multimodal model was created by concatenating tabular features with image embeddings. A Ridge Regression model was then trained on this combined feature set. Ridge regression was chosen to handle the increased dimensionality and reduce overfitting.

Only samples that had both tabular data and satellite images were used for multimodal training and evaluation.

Results:

The performance of the tabular-only baseline model was compared with the multimodal model.

The multimodal approach showed better predictive performance, indicating that satellite images provide useful additional information for property valuation. Even a modest improvement in error metrics demonstrates the value of incorporating visual neighbourhood context.

```
MULTIMODAL MODEL PERFORMANCE (TRAIN SUBSET)
RMSE: 0.2210
MAE : 0.1727
```

Explainability:

Although explicit Grad-CAM visualizations were not generated, the multimodal framework allows for interpretability through CNN-based feature extraction. Conceptually, Grad-CAM could highlight areas such as green spaces, water bodies, or road connectivity that contribute positively to price predictions.

This aligns well with real-world intuition about what makes a neighbourhood desirable.

Limitations:

- Satellite images were used for a subset of the dataset due to API and computational constraints.
- The CNN was used as a feature extractor without fine-tuning.
- Multimodal evaluation was performed on aligned subsets rather than the full dataset.

These limitations do not affect the core conclusion of the project.

Conclusion:

This project demonstrates that combining tabular housing data with satellite imagery leads to better property price prediction. The multimodal regression framework successfully captures both numerical and visual aspects of real estate valuation.

The results highlight the importance of neighbourhood environment in determining property value and show how multimodal learning can be applied to real-world financial and spatial problems.