REPORT

Title: *Satellite Imagery-Based Property Valuation*

Enroll num: 24114065 **Name**:Pavan Kumar

Overview about the Project:

This project implements a **Multimodal Regression Pipeline** that integrates structured tabular data with satellite imagery to enhance the accuracy of real estate valuations. By programmatically capturing environmental features like green space and road density, the system quantifies "curb appeal" and neighborhood context alongside traditional financial metrics. The final model employs a **data-fusion architecture** to bridge the gap between numerical trends and visual aesthetics, providing a comprehensive valuation framework for modern analytics.

Final Goal :

The goal of this project is to build a **multimodal regression system** that enhances property valuation by fusing traditional tabular data with satellite-derived environmental insights to capture the impact of neighborhood context on market value.

Project Approach :

- **Dual-Stream Data Pipeline and data fusion :** Synchronized structured housing records with geospatial coordinates to programmatically fetch neighborhood satellite imagery via the **Mapbox API**.
- **Feature Extraction:** Utilized a pretrained **ResNet-18** CNN to convert raw visual data into 512-dimensional embeddings, capturing "curb appeal" and environmental context numerically.
- **Multimodal Fusion:** Implemented a **Late Fusion** strategy by concatenating tabular features with visual embeddings, using **XGBoost** to learn the complex relationship between physical house attributes and their surroundings.
- **Visual Explainability:** Employed **Grad-CAM** to highlight specific image regions (like green spaces or roads) that influenced the valuation, ensuring the model's predictions are interpretable.

Data Fetcher :

Data Acquisition Approach

- **API Integration:** Used the Mapbox Static Images API to programmatically convert GPS coordinates (latitude/longitude) into high-resolution satellite tiles.

- **Standardized Preprocessing:** Fetched images at 224x224 resolution with Zoom Level 18, ensuring visual consistency and compatibility with the ResNet-18 model.
- **Automated Pipeline:** Developed a robust Python script with rate-limiting (0.2s delay) and checkpointing to efficiently download thousands of images without data loss or API bans.
- **Geospatial Linking:** Mapped each image to its corresponding tabular record using unique Property IDs to ensure perfect alignment for the multimodal model.

Data Loading and Description:

The data loading and description phase involves ingesting the **King County Housing dataset**, which consists of 21 structural and geospatial attributes. The process uses **Pandas** to load training and testing subsets, specifically targeting features like square footage, grade, and year built. Crucially, it leverages the **latitude and longitude** coordinates to serve as the link between these numerical records and their corresponding satellite imagery, ensuring a synchronized dataset for multimodal analysis.

Data Cleaning and Missing Value Handling :

To address data quality, I performed **missing value analysis** to ensure a complete dataset; any null entries were resolved via **median imputation** to maintain statistical distribution without outlier bias.
Simultaneously, I performed **feature pruning** by removing non-predictive columns like `id` and `date`, which could lead to overfitting. This streamlined the data into a high-signal matrix focused purely on structural and geospatial attributes.

Exploratory Data Analysis: Visualization of Risk Drivers

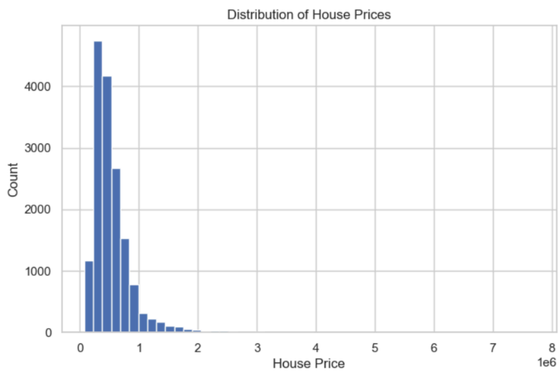
In this section, we perform an in-depth exploratory data analysis (EDA) using rich visualizations to uncover key patterns and relationships relevant to credit card default risk.

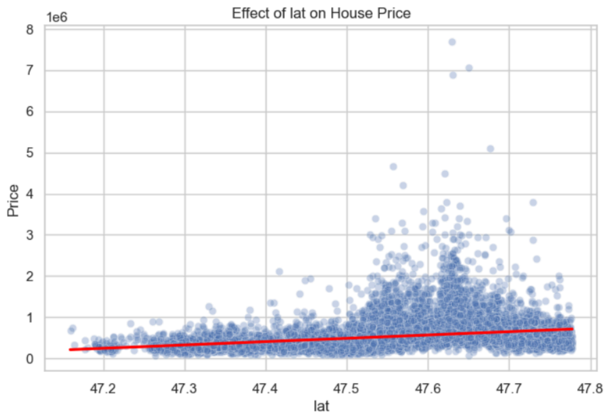
Scaling and Splitting :

Train-Test Split: Partitioned the dataset into training (80%) and testing (20%) subsets to validate model generalizability and prevent overfitting on unseen data.
Standard Scaling: Applied **StandardScaler** to numerical features, centering data around a zero mean and unit variance.
Optimization: Scaling ensures that high-magnitude features (like `sqft_living`) do not numerically overwhelm smaller variables (like `bedrooms`), facilitating faster convergence for **XGBoost**.
Leakage Prevention: Fit the scaler only on the training data and transformed the test data to maintain strict evaluation integrity.

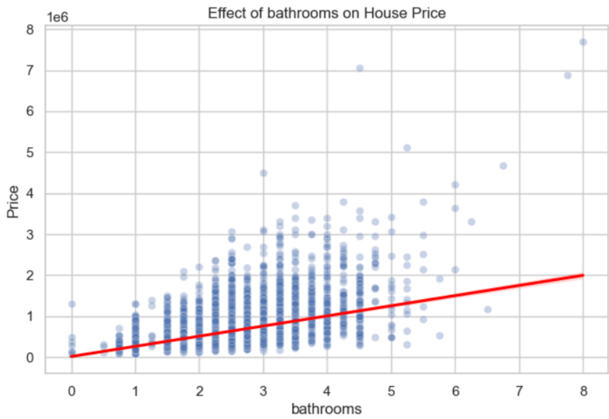
<u>Exploratory Data Analysis</u> :

How Features Affect House Prices

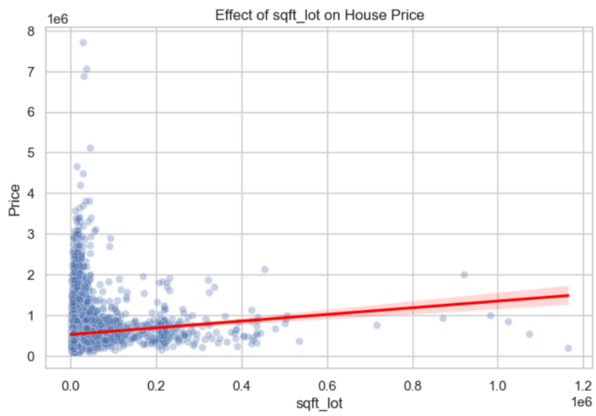




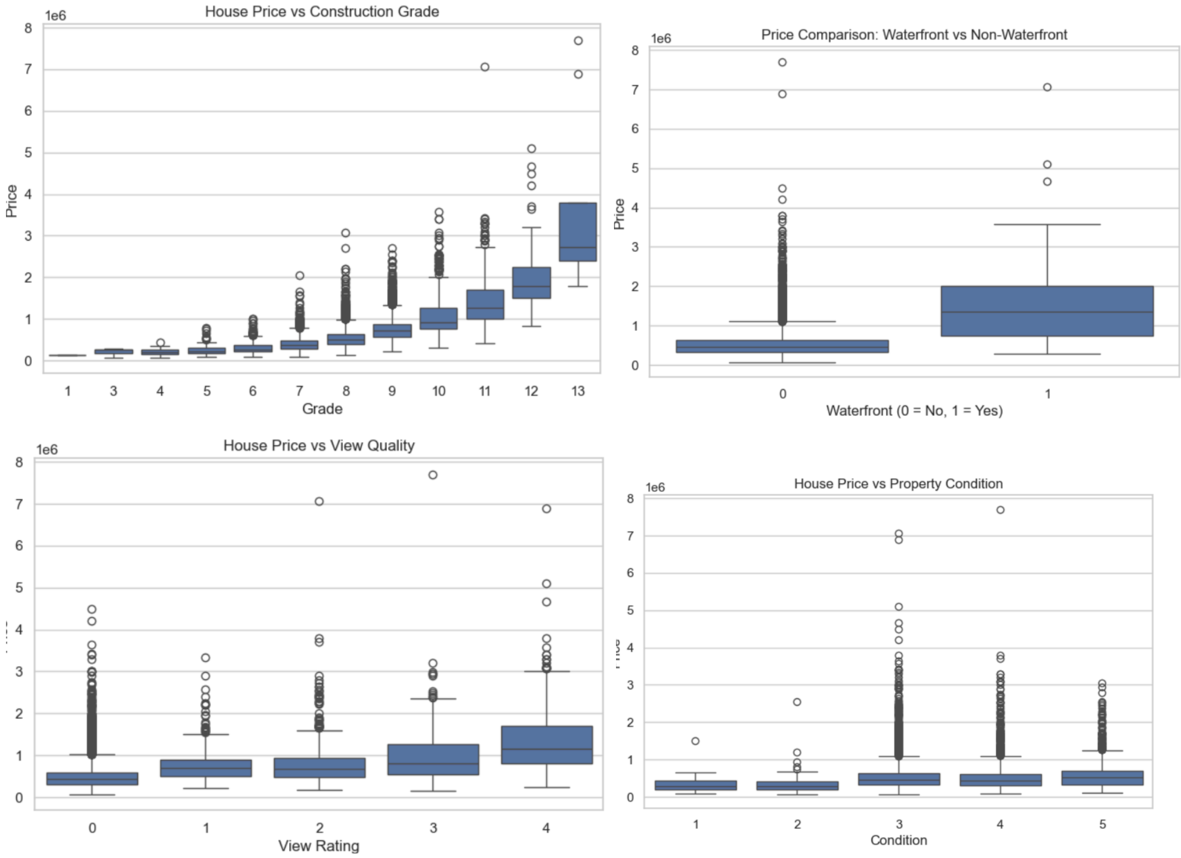
The above 2 graphs shows how House prices varies with count and lat



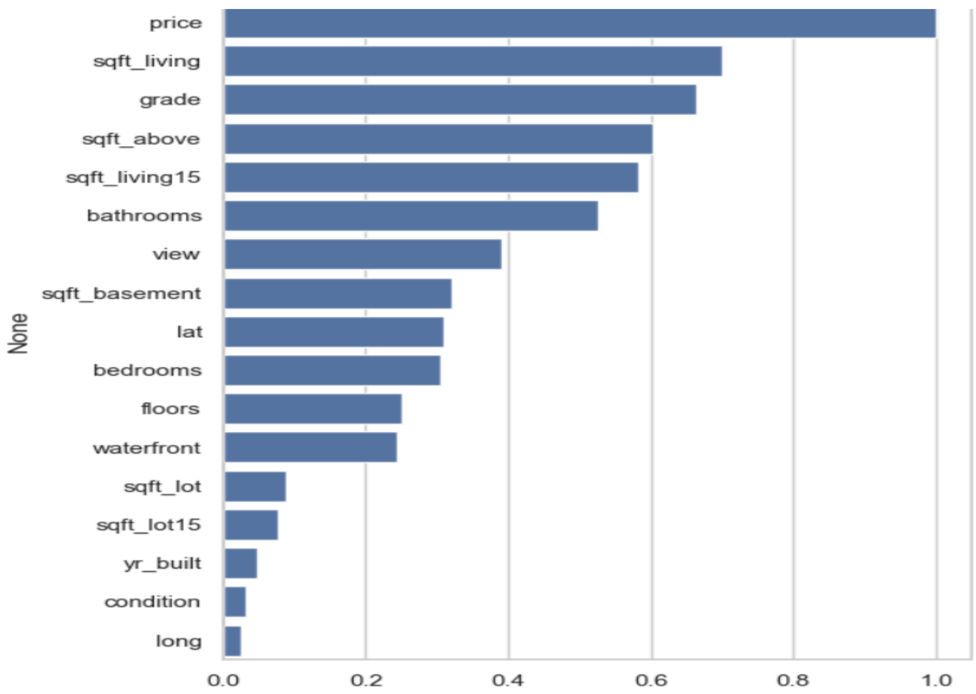
The above 2 graphs shows how House prices varies with sqft_living and bathrooms



The above 2 graphs shows how House prices varies with sqft_lot and bedrooms



The above 4 graphs shows us how House Price varies with Construction Grade , Water front , Quality and Condition.



The above graph shows the Correlation of Features with House prices.:

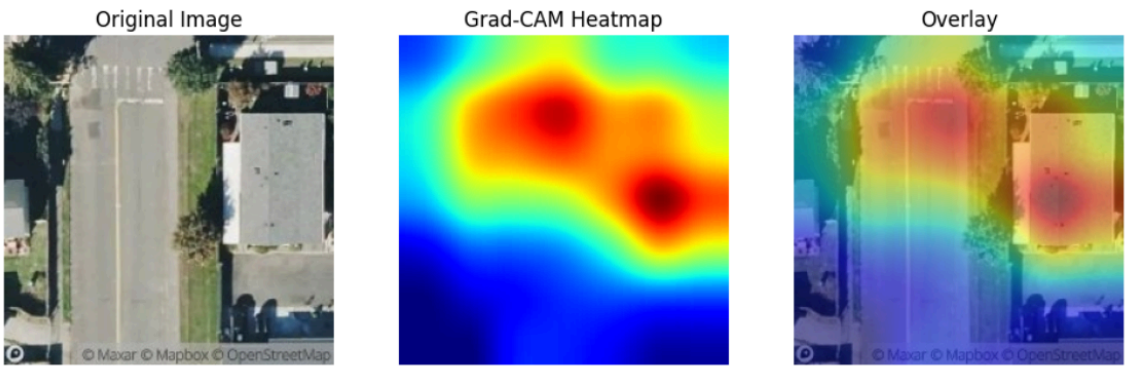
Key Insights from Feature Analysis

- sqft_living is the strongest driver of house price.
- grade has a larger impact than the number of bedrooms.
- Waterfront properties show a significant price premium.
- Bedrooms alone have weak explanatory power without size/quality.
- XGBoost relies heavily on construction quality, size, and location.

These insights justify the use of non-linear, tree-based models and motivate the inclusion of visual (satellite) features in the multimodal model.

Grad-CAM Explainability for Satellite Imagery :

I implemented **Grad-CAM** (Gradient-weighted Class Activation Mapping) to provide visual explainability for the model's price predictions. By calculating the gradients of the target property value with respect to the final convolutional layer of the **ResNet-18**, the system generates a heatmap that highlights specific areas within the satellite imagery—such as green cover or road density—that most heavily influence the valuation. This process effectively converts the deep learning model from a "black box" into an interpretable tool, allowing stakeholders to see exactly which neighborhood features are driving the asset's market value.



Model Evaluation:(Model with its R2 score and Rmse score)

1. For Image only :

```
Image-only RMSE: 318995.31241696957
Image-only R²: 0.1433895230293274
```

2.For Tabular Only:

	Model	RMSE	R2
0	Linear Regression	193064.864288	0.702969
1	Random Forest	130138.003233	0.865040
2	XGBoost	120336.862648	0.884603

3.For MultiModal Only :

```
Multimodal RMSE: 115952.7230555626
Multimodal R²: 0.8868181705474854
```

Models Comparision :

	Model	RMSE	R2
0	Tabular Only	120336.862648	0.884603
1	Image Only	318995.312417	0.143390
2	Tabular + Image (Multimodal)	115952.723056	0.886818

Best R2 score : 0.886818
Best RMSE score: 115,952.723

Final step:

Finalizing the project, the **Multimodal Model** outperformed standalone tabular approaches, yielding a higher **R2** and lower **RMSE**. This confirms that satellite imagery provides vital environmental context—like green cover and density—that numerical data lacks. Consequently, this hybrid architecture was used to generate the final **submission CSV**, containing only the id and predicted_price for each asset.

Why Multimodal?

Based on the performance metrics, the **Multimodal Model** is clearly the superior architecture for property valuation. By achieving a higher **R^2score** and a lower **RMSE** compared to tabular-only baselines, the model proves that satellite imagery captures critical "hidden" variables—such as neighborhood prestige and environmental quality—that numerical data alone cannot represent.

The integration of visual embeddings effectively reduces prediction error and provides a more comprehensive understanding of market value, making it the finalized choice for the project's predictive pipeline.