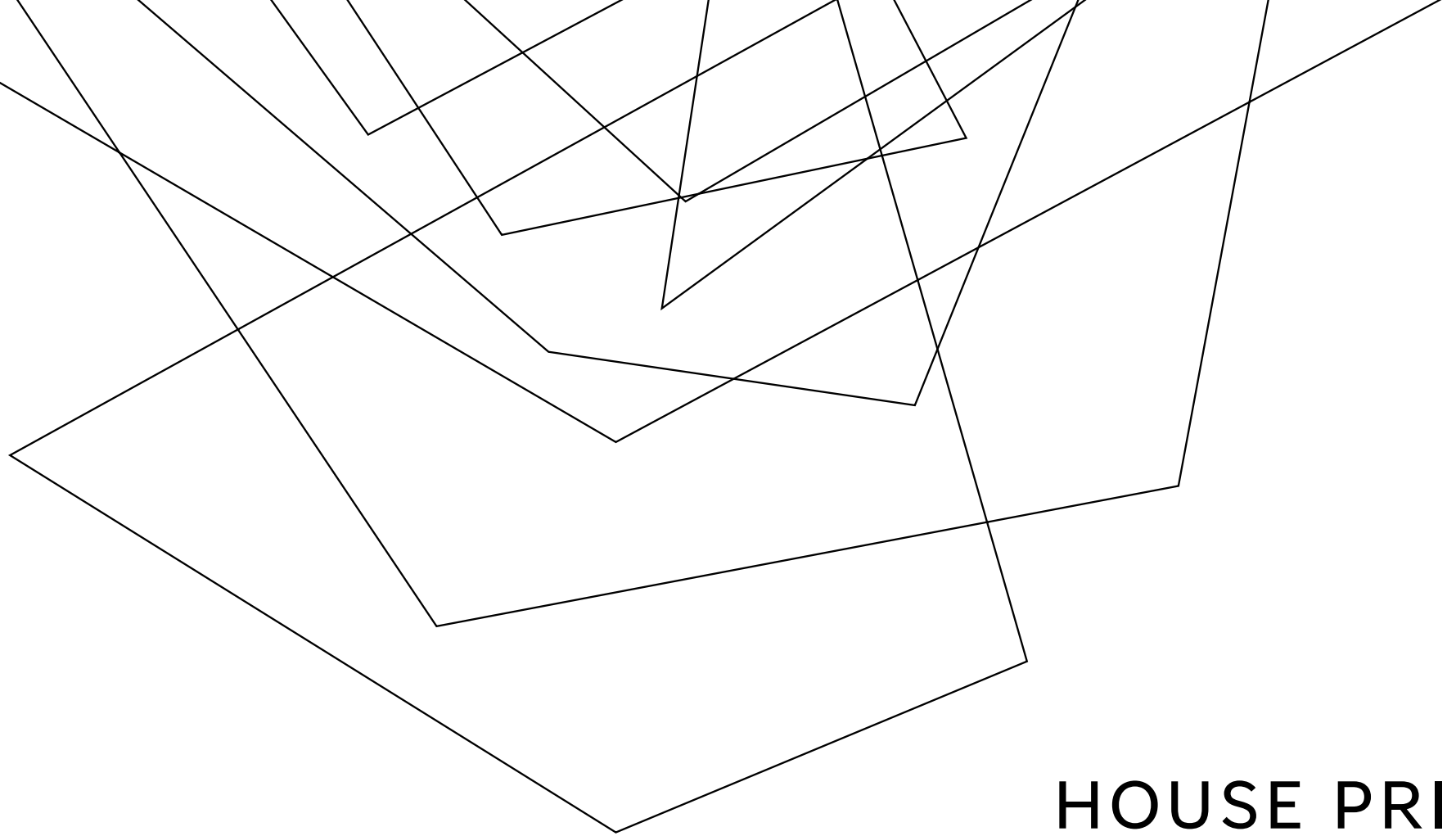


HOUSE PRICE PREDICTION

AGENDA

Introduction

Problem Statement

Data Description

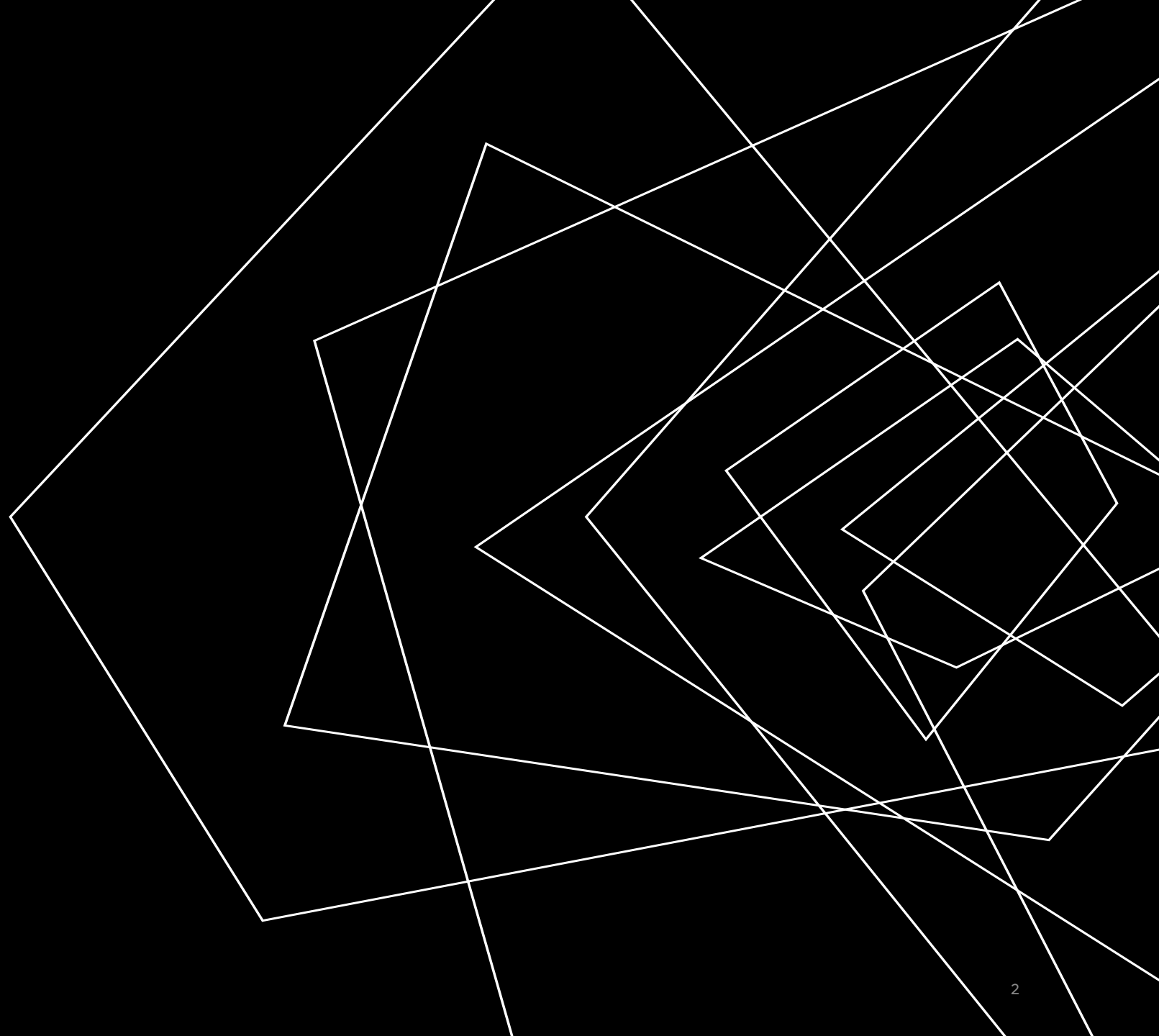
Methodology

Data Preprocessing

Model Development

User Interface

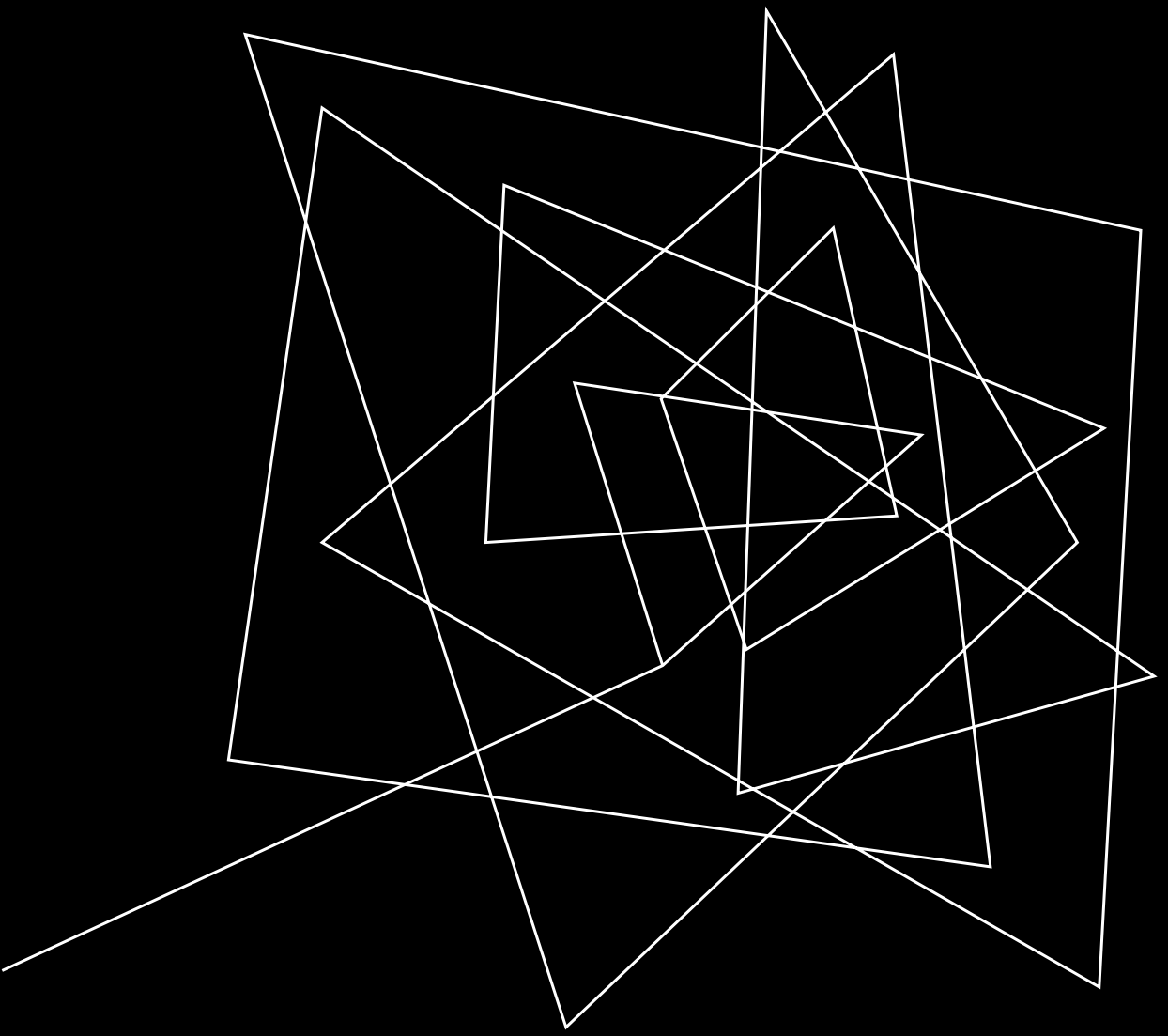
Results and Analysis



INTRODUCTION

This project aims to analyze residential property prices to uncover insights into the factors that influence the housing market. By exploring various variables related to property characteristics, location, market trends, and economic indicators, we seek to provide a comprehensive understanding of the dynamics driving home prices.

- **Key Goals:**
  - **Identify Key Factors:** Determine the primary elements that affect property prices, including size, condition, and location.
  - **Uncover Trends:** Recognize patterns and trends in the housing market that impact property values.
  - **Inform Decision-Making:** Provide valuable insights for real estate professionals, investors, policymakers, and prospective homeowners to guide their decisions regarding property investments, pricing strategies, and policy development.



PREDICTIVE POWER UNLEASHED

PROBLEM STATEMENT

The real estate market is influenced by a complex interplay of factors, making it challenging to accurately assess residential property values and predict market trends. Stakeholders such as real estate professionals, investors, policymakers, and prospective homeowners often struggle to understand the underlying dynamics that drive housing prices.

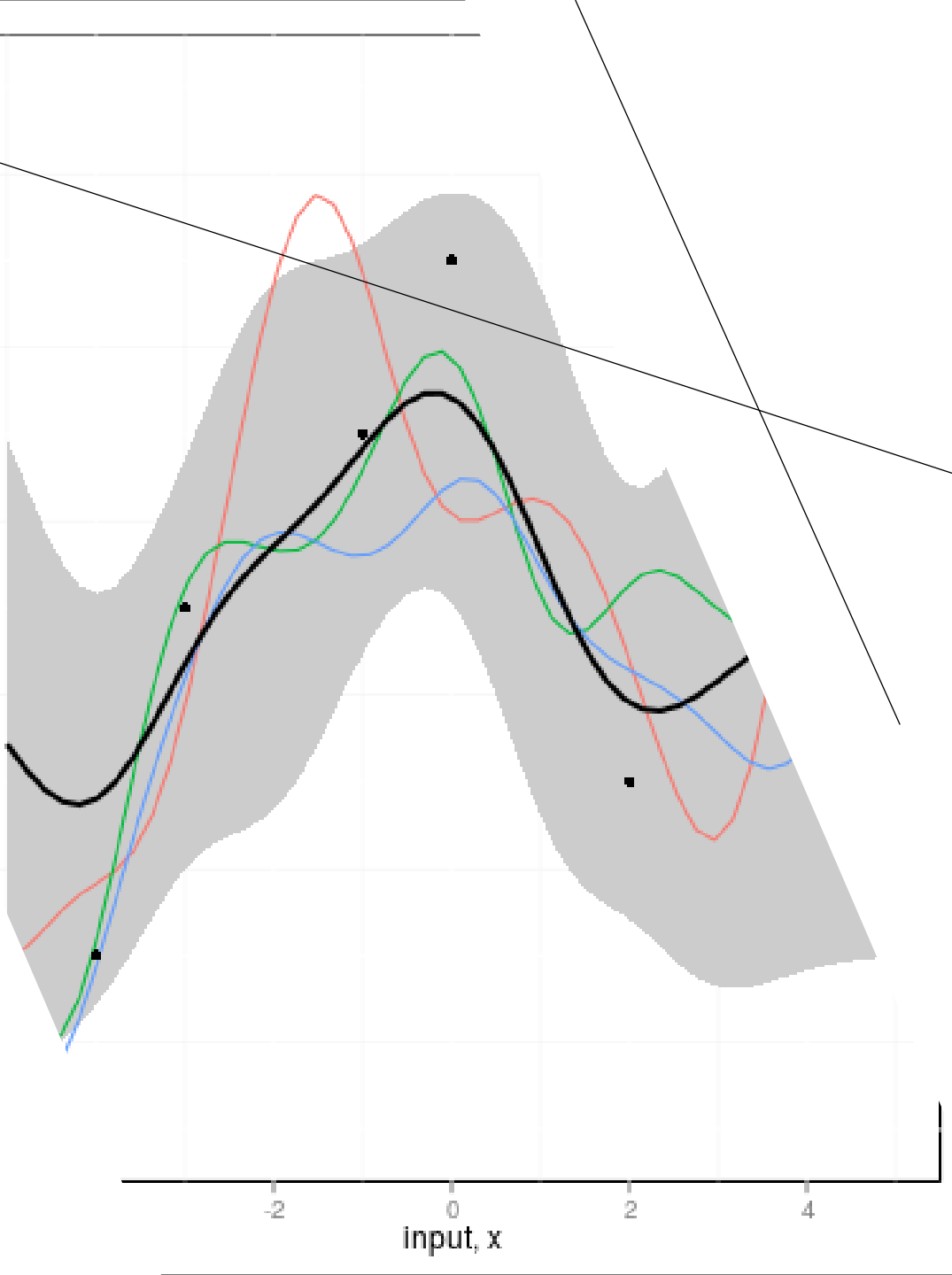
The challenge lies in identifying and quantifying the impact of various property characteristics, location attributes, market trends, and economic indicators on residential property prices. Without a clear understanding of these factors, making informed decisions regarding real estate investments, pricing strategies, and policy development becomes difficult.

This project aims to address these challenges by analyzing a comprehensive dataset of residential properties, exploring key variables, and uncovering the correlations and trends that influence housing market dynamics.

DATA DESCRIPTION

The dataset encompasses various attributes of residential properties, including price, size (square footage of living space and lot), and the number of bedrooms and bathrooms. It also includes features like the number of floors, property condition, and location details (city, state, ZIP code). Special attributes such as waterfront views, the year built, and renovations provide additional context. This data facilitates a comprehensive analysis of the factors influencing housing prices.

price	bedrooms	bathrooms	sqft_living	sqft_lot	floors	waterfront	view	condition	sqft_baseme	yr_built	yr_renovatec	statezip	country
313000	3	1.5	1340	7912	1.5	0	0	3	0	1955	2005	WA 98133	USA
2384000	5	2.5	3650	9050	2	0	4	5	280	1921	0	WA 98119	USA
342000	3	2	1930	11947	1	0	0	4	0	1966	0	WA 98042	USA

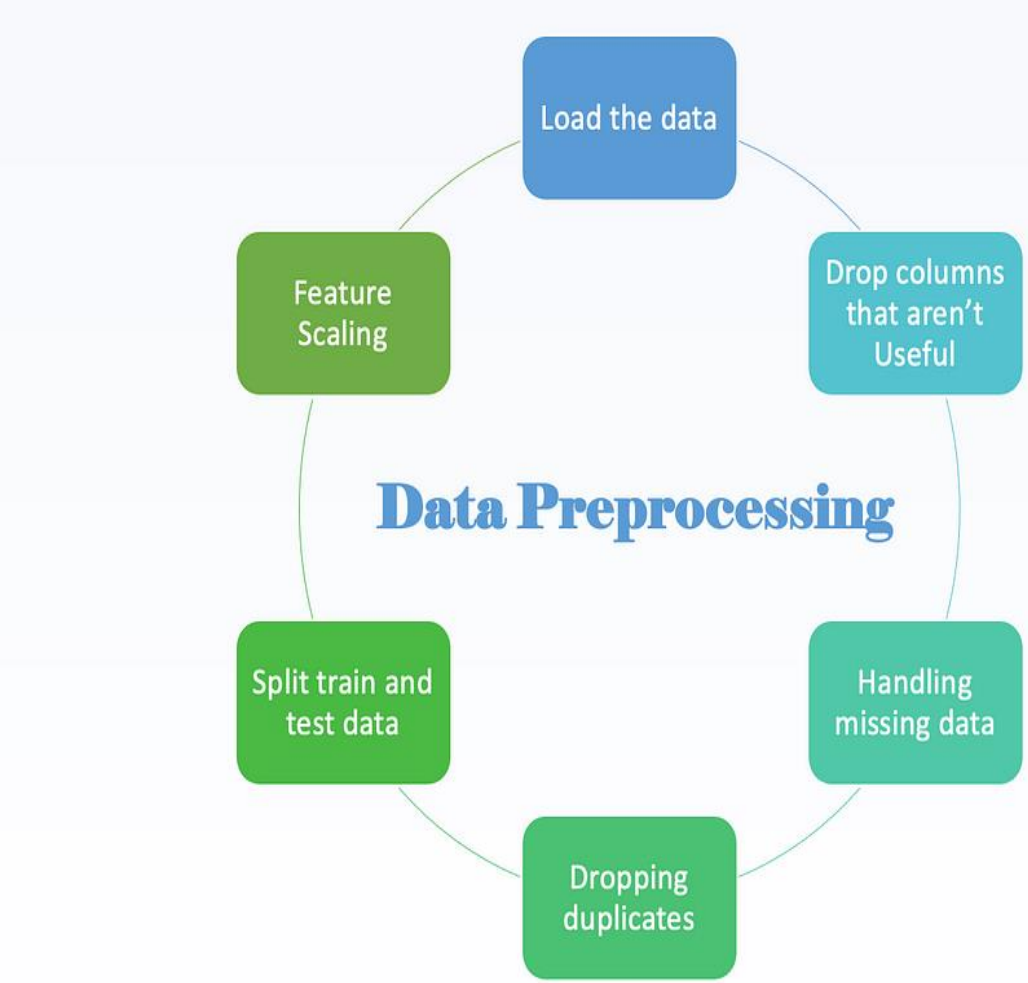


METHODOLOGY

The analysis involved collecting and preprocessing a dataset of residential properties, including cleaning data and engineering new features. Exploratory Data Analysis (EDA) used visualizations and summary statistics to identify key relationships between features. Linear regression was applied to model the impact of selected features on property prices. The model's performance was evaluated using metrics like Mean Squared Error (MSE) and R-squared, ensuring accuracy through cross-validation. Insights from the model informed recommendations for stakeholders to support data-driven decisions in the housing market.

DATA PREPROCESSING

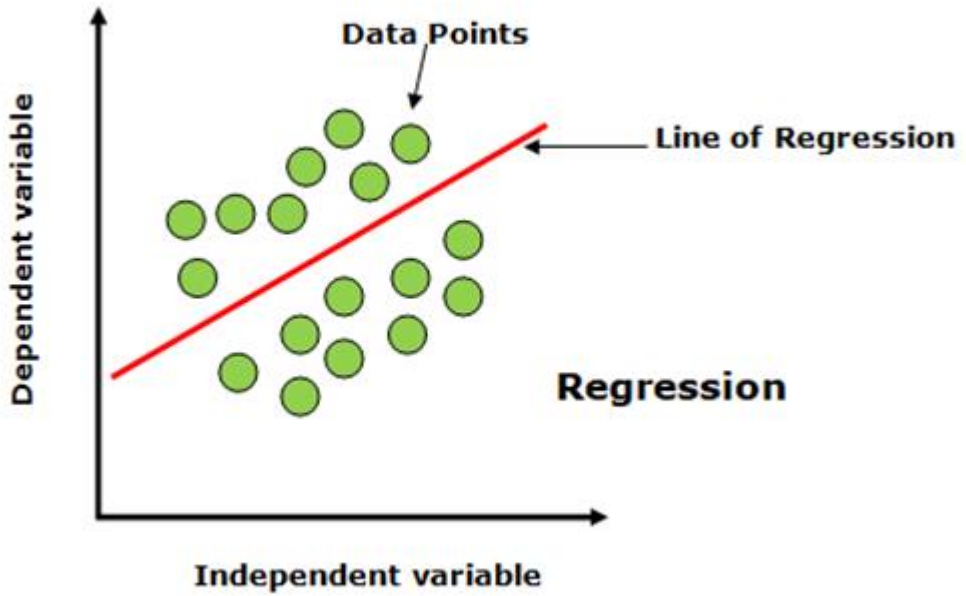
Data preprocessing involved several key steps to prepare the dataset for analysis. First, the dataset was cleaned to handle missing values, correct inconsistencies, and remove any duplicate entries. Next, data transformation was performed, including normalization or scaling where necessary to ensure consistent measurement units. Feature engineering was used to create new variables that enhance analysis, such as calculating property age or total rooms. Finally, outliers were identified and addressed to prevent their distortion of the analysis. These preprocessing steps ensured that the data was accurate, consistent, and ready for modeling.



MODEL DEVELOPMENT

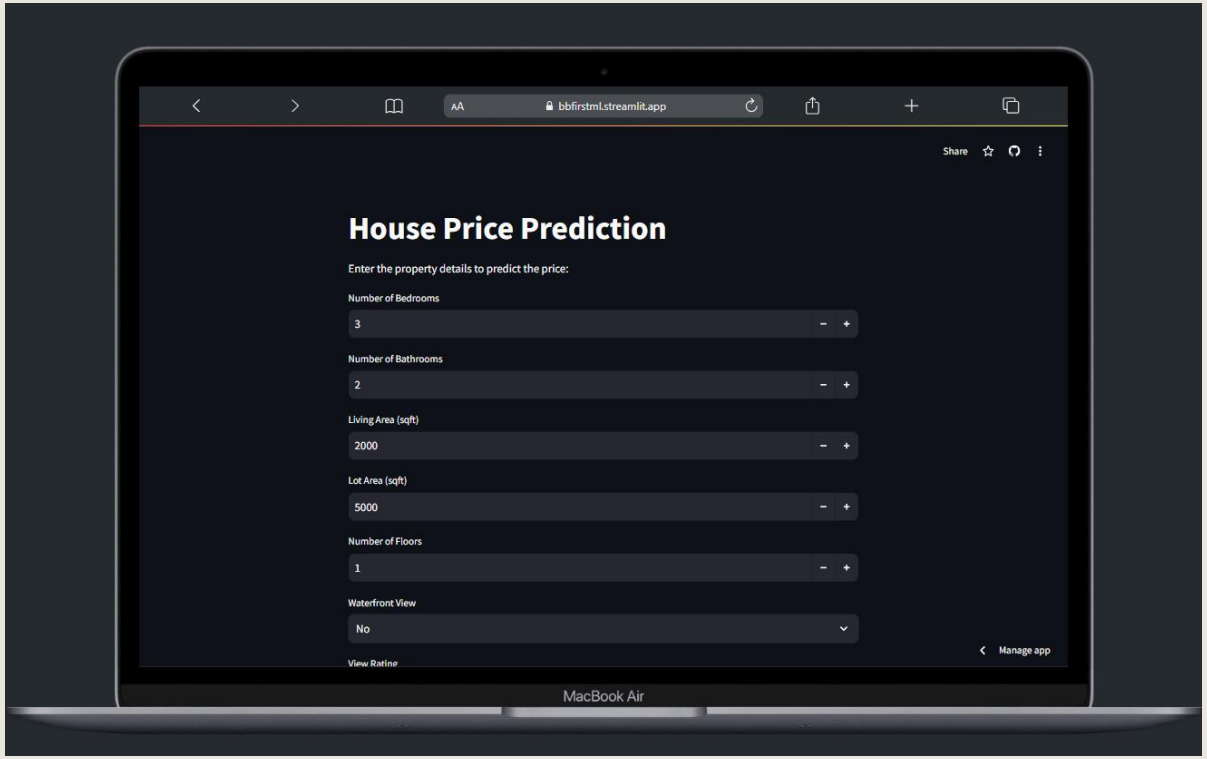
For the model development, linear regression was chosen to analyze the relationship between various property features and their sale prices. This method was selected due to its effectiveness in understanding and quantifying the impact of multiple independent variables on a continuous dependent variable.

The process began with preparing the dataset by selecting relevant features and splitting it into training and testing sets. The linear regression model was then trained using the training data, learning how different features like square footage, number of bedrooms, and property condition influence the price.



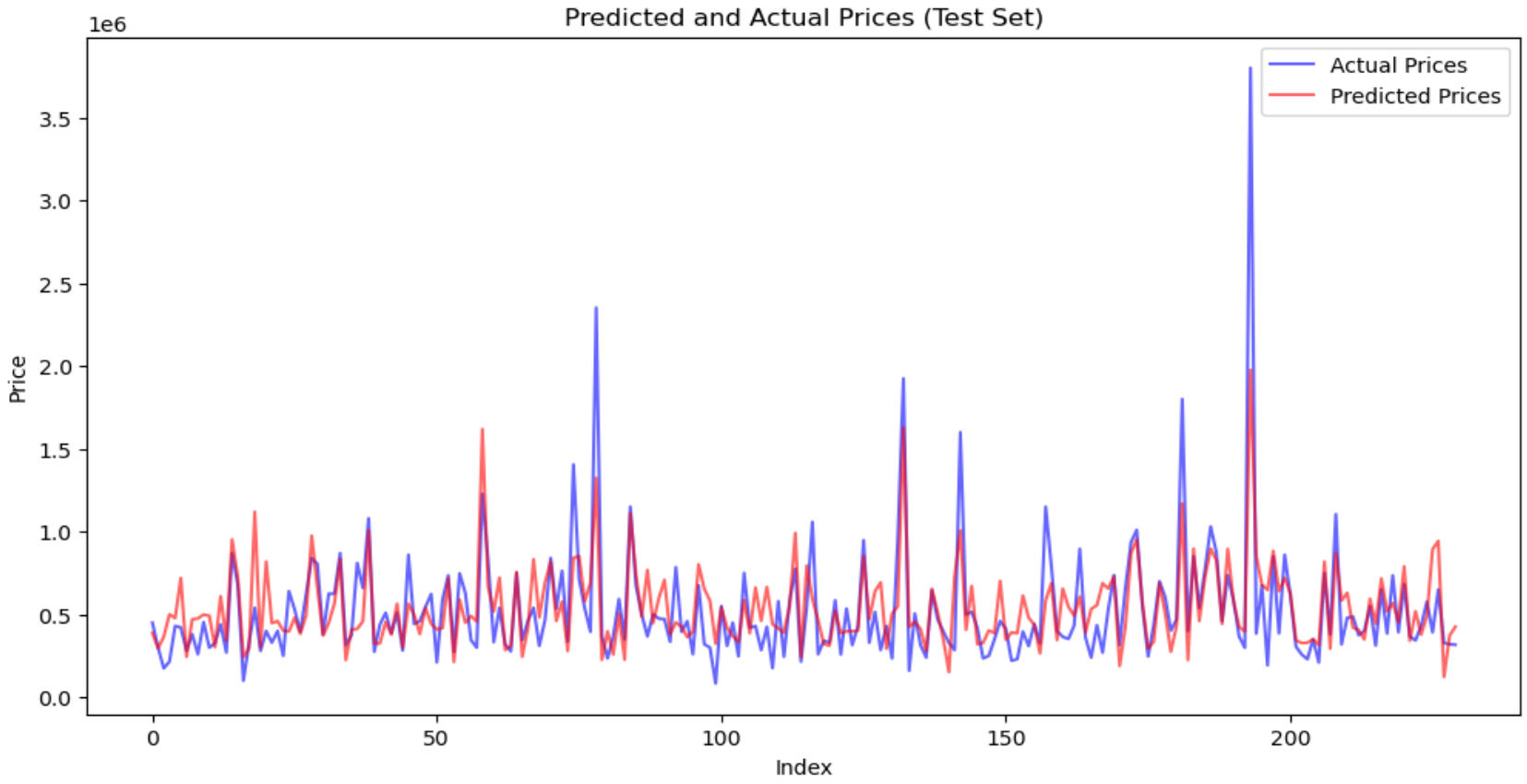
USER INTERFACE

- The user interface (UI) was developed using Streamlit, providing a straightforward platform for interacting with the property price analysis model. The UI features:
- **Input Forms:** Allows users to enter property details such as square footage, number of bedrooms, bathrooms, and location attributes.
- **Submit Button:** Processes the entered data to generate predictions from the model.
- **Results Display:** Shows the predicted property prices and related outputs based on the model's calculations.



RESULT'S & ANALYSIS

The analysis utilized linear regression to model and predict residential property prices. One of the key visualizations is the **Actual vs. Predicted Price** graph, which compares the model's predictions with the actual sale prices of properties.



THANK YOU

Bhargav Bhatt

9426027083

bhargavmbhatt24@gmail.com

Bbfirstml.streamlit.app