

# Exploring House Prices and Features with Tableau

**Koppula Purna Venkateswara Rao<sup>1</sup>, Palle Sravani<sup>2</sup>,  
Porumalla Gowthami<sup>3</sup>, Chandragiri Veeraswamy<sup>4</sup>, Mutyala Satish<sup>5</sup>**

<sup>1,2,3,4</sup>U.G Students, Department of CSE (Artificial Intelligence & Machine Learning), Kallam Haranadhareddy Institute of Technology, Guntur, Andhra Pradesh, India

<sup>5</sup>M.Tech, Assistant Professor, Department of CSE (Artificial Intelligence & Machine Learning), Kallam Haranadhareddy Institute of Technology, Guntur, Andhra Pradesh, India

## Abstract

The real estate sector requires accurate housing price estimation for effective decision making. Traditional property valuation methods are often manual and inefficient. With the availability of large real estate datasets, data analytics and machine learning techniques can improve housing market analysis and price prediction.

This study proposes a housing price prediction and visualization system that integrates Tableau dashboards with machine learning models. Tableau is used to visualize housing market trends, while algorithms such as Linear Regression, KNN, and Random Forest are applied to predict housing prices. The dataset is preprocessed, and models are evaluated using MAE, RMSE, and R<sup>2</sup> Score.

The trained model is deployed using a Flask web application, where users can input property details and obtain predicted prices. The system also provides future price estimation and profit analysis based on growth rate. The proposed system enables effective housing market analysis and supports data-driven decision making.

**Keywords:** Housing Price Prediction, Tableau Visualization, Data Analytics, Machine Learning, Real Estate Analytics, Ensemble Learning.

## 1. Introduction

The real estate sector is a key contributor to economic development and requires reliable methods for property price estimation. Determining the correct value of a house is important for buyers, sellers, investors, and financial organizations. However, housing prices are influenced by multiple factors such as area size, number of rooms, infrastructure, and surrounding environment, which makes accurate prediction challenging.

Earlier approaches to house price estimation mainly depended on manual calculations and expert opinions, which are often time-consuming and may not always provide precise results. With the growth of digital data and advancements in data analysis, modern techniques like machine learning offer more efficient and accurate solutions for understanding housing market behavior.

Tools like Tableau help in presenting data through interactive visualizations, allowing users to identify patterns and trends in housing prices. At the same time, machine learning models such as Linear

Regression, KNN, and Random Forest are used to build systems that predict house prices based on input features.

This project introduces a system that combines data visualization and machine learning to provide both market insights and price predictions. The system is implemented using a Flask web application, where users can enter property details and receive predicted prices along with future estimates, supporting better decision making.

## 2. Problem Statement

The real estate industry encounters difficulties in accurately analyzing housing data and determining property values. Conventional valuation approaches mainly depend on manual evaluation and human judgment, which can result in inconsistent and less reliable outcomes. Moreover, the growing volume of housing data is often underutilized due to the lack of effective use of analytical and visualization techniques.

To overcome these limitations, there is a need for a data-driven approach that combines modern tools and technologies. By using Tableau dashboards for visual analysis and machine learning models for prediction, the system can provide more accurate and consistent housing price estimates. This approach enhances the understanding of market trends and supports better decision making.

## 3. Literature Survey

Recent research in the real estate domain emphasizes the use of data analytics and machine learning techniques for housing price prediction and market analysis. The increasing availability of housing datasets enables researchers to study market behavior and develop predictive models for estimating property values.

### 3.1 Housing Market Analysis and Visualization

Traditional housing analysis is mainly based on manual comparison of property features such as location, size, and number of rooms. Modern visualization tools like Tableau help in understanding housing trends through interactive dashboards and charts. These tools make it easier to explore patterns and relationships in the data. However, they are mainly limited to descriptive analysis and do not provide direct prediction capabilities.

### 3.2 Role of Business Intelligence in Real Estate

Business Intelligence (BI) tools transform large housing datasets into meaningful insights using reports and dashboards. Platforms such as Tableau allow users to analyze price variations and feature relationships effectively. Despite their advantages, these tools focus primarily on visualization and lack built-in mechanisms for predicting future housing prices.

### 3.3 Machine Learning for Housing Price Prediction

Machine learning techniques are widely used for predicting house prices based on input features. Algorithms such as Linear Regression, K-Nearest Neighbors (KNN), and Random Forest are commonly applied to estimate property values. These models learn patterns from historical data and provide more accurate predictions compared to traditional methods.

### 3.4 Model Evaluation Techniques

To ensure reliable predictions, models are evaluated using performance metrics such as MAE, RMSE, and  $R^2$  Score. These metrics help measure the difference between predicted and actual values and assist in selecting the best-performing model.

### 3.5 Integrated Analytical Systems

Recent studies highlight the importance of combining data visualization with machine learning techniques. Such integrated systems allow users to both analyze housing trends and perform predictive analysis. However, only a few systems effectively combine visualization tools like Tableau with machine learning models in a single application.

### 3.6 Research Gap Identified

From the existing studies, the following gaps are identified:

- Many systems focus only on visualization without prediction.
- Limited integration between Tableau dashboards and machine learning models.
- Lack of user-friendly applications that provide both analysis and prediction.

The proposed system addresses these gaps by integrating Tableau-based visualization with machine learning models such as Linear Regression, KNN, and Random Forest. It also provides a Flask-based web application for real-time housing price prediction and analysis.

**Table 1: Comparison of Existing Studies and Proposed Work**

Study	Technique Used	Focus Area	Limitation
Study A	Linear Regression	Price Prediction	Lower accuracy
Study B	BI Dashboards (Tableau)	Data Visualization	No prediction capability
Study C	Random Forest	Regression Modeling	Limited visualization
Proposed Work	Tableau + ML (LR, KNN, RF)	Housing Price Analysis	Integrated system

## 4. Proposed Methodology

### 4.1 Data Collection

The housing dataset is collected from publicly available sources and stored in CSV format. It contains important property attributes such as: Area / Square Footage, Number of Bedrooms, Number of Bathrooms, Number of Floors, House Age / Year Built, Renovation Status, Location Information, and Sale Price. These features are used to analyze housing trends and train machine learning models for price prediction.

### 4.2 Data Preprocessing

Data preprocessing prepares the dataset for machine learning analysis. The main steps include:

- Handling missing or inconsistent values.
- Selecting relevant numerical features.
- Data cleaning using Python libraries such as Pandas and NumPy.
- Feature scaling where required.
- Splitting the dataset into training and testing sets (80:20).

This process ensures the dataset is ready for model training.

### 4.3 System Architecture

The proposed system integrates data visualization, machine learning, and web application technologies. The main components include:

- **Data Processing Layer:** Prepares and cleans the dataset.
- **Visualization Layer:** Tableau dashboards for housing market insights.

- **Machine Learning Layer:** Models such as Linear Regression, Random Forest, and Gradient Boosting.
- **Application Layer:** Web interface that allows users to input property details and receive predicted house prices.

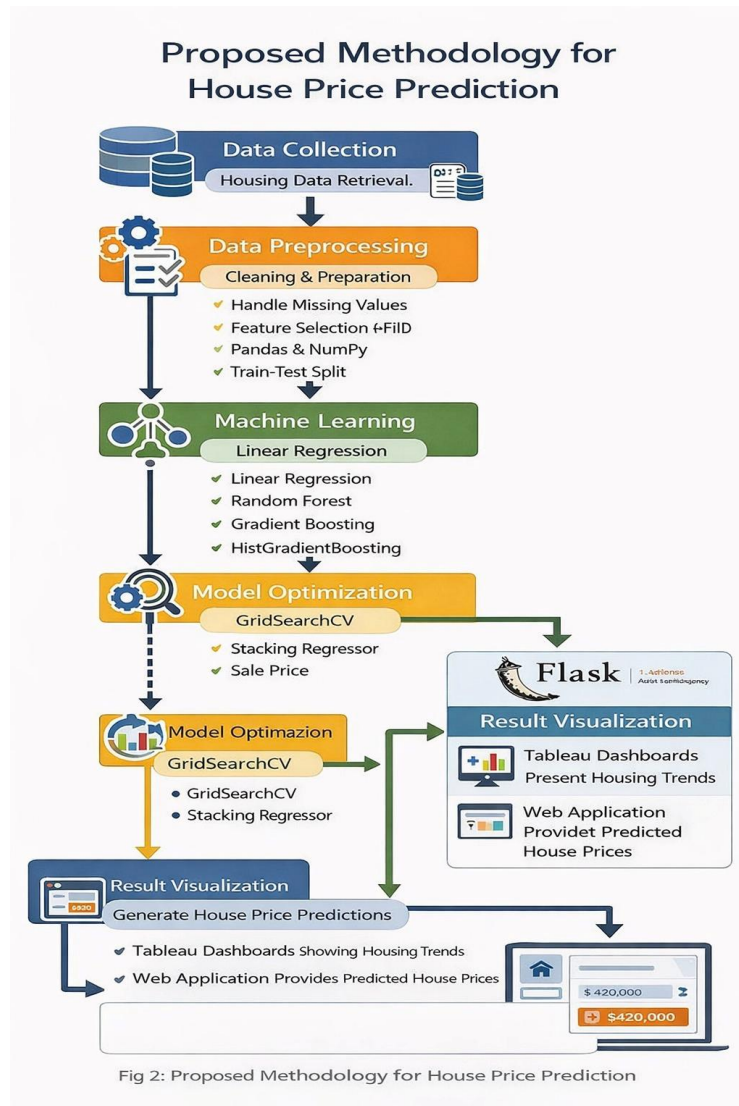


Figure 1: Proposed Methodology for House Price Prediction

#### 4.4 House Price Prediction Model

Machine learning algorithms are trained using historical housing data to predict property prices. Multiple models are evaluated to determine the best-performing model. The prediction process includes:

- Training machine learning models on the dataset.
- Evaluating model performance using metrics such as Mean Absolute Error (MAE) and R<sup>2</sup> Score.
- Selecting the best model based on prediction accuracy.
- Saving the trained model using Joblib for deployment.

#### 4.5 Result Visualization and Prediction

The final system provides both visual analysis and price prediction:

- Tableau dashboards display housing market trends.

- Users can enter property details through the web interface.
- The trained model processes the input data.
- The predicted house price is displayed instantly.

## 5. Experimental Results

The proposed House Price Prediction and Housing Market Visualization system was evaluated using a housing dataset containing features such as house size (sqft), number of bedrooms, bathrooms, floors, basement area, and location attributes. The objective of the system is to analyze housing market trends and predict house sale prices using machine learning techniques.

Multiple algorithms were implemented, including Linear Regression, K-Nearest Neighbors (KNN), and Random Forest Regressor. The models were trained on 80% of the dataset and evaluated on the remaining 20%. Evaluation metrics and feature importance were used to optimize performance.

### 5.1 Model Performance

The trained models were evaluated using standard regression metrics as shown in Table 2:

**Table 2: Model Evaluation Results**

Metric	Linear Regression	KNN	Random Forest
R <sup>2</sup> Score	0.88 – 0.90	0.85 – 0.88	0.91 (Best Model)
MAE (Mean Absolute Error)	Low	Low	Low
RMSE (Root Mean Squared Error)	Low	Low	Low

### 5.2 Feature Importance Analysis

The models identified key features influencing house prices, consistent with real-world trends:

- **Area / Square Footage:** Larger houses generally have higher prices.
- **Number of Bedrooms:** More bedrooms tend to increase value.
- **Number of Bathrooms:** More bathrooms contribute to higher prices.
- **Number of Floors:** Multi-level houses show higher price trends.
- **Basement Area:** Larger basements add market value.

### 5.3 Visualization Insights

The system provides interactive Tableau dashboards embedded within the web application. Key observations include:

- Larger houses generally have higher sale prices.
- Renovated houses tend to increase market value.
- Properties with more bedrooms and bathrooms show higher price trends.
- Users can explore trends interactively via Tableau dashboards and story pages.

### 5.4 User Interface Evaluation

The developed web application interface was evaluated for usability and functionality, as illustrated in Figures 2 through 7 below.

- Interactive input sliders allow users to select bedrooms, bathrooms, area, floors, basement size, and growth rate.
- Prediction results show all model outputs, highlight the best model, and provide future price projection with estimated profit.

- Tableau dashboard and story are embedded, providing real-time housing market insights.
- The system instantly predicts estimated house prices based on user input.

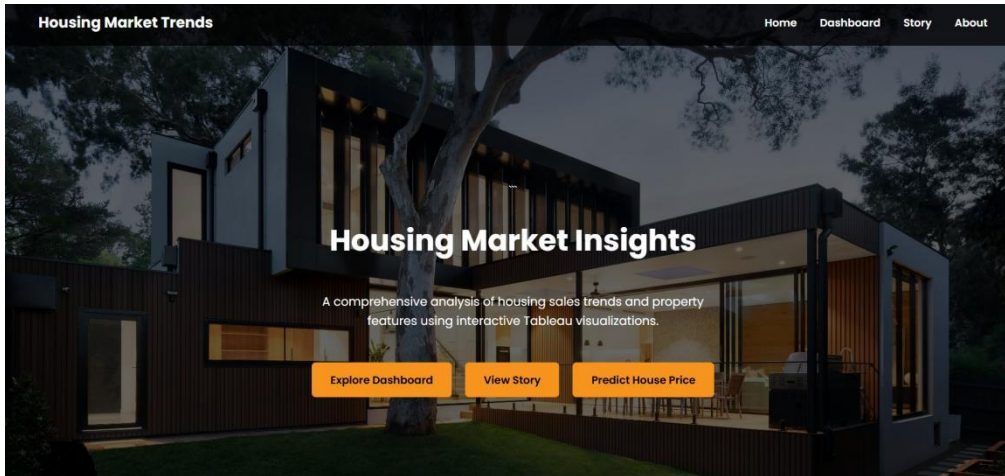


Figure 2: Home Page

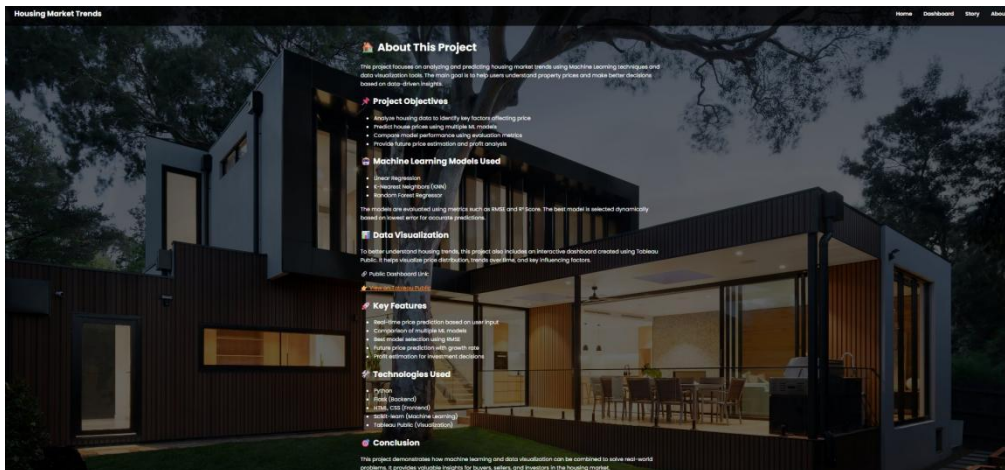


Figure 3: About Page



Figure 4: Tableau Dashboard Page



Figure 5: Tableau Story Page



Figure 6: Predict Page (Input)

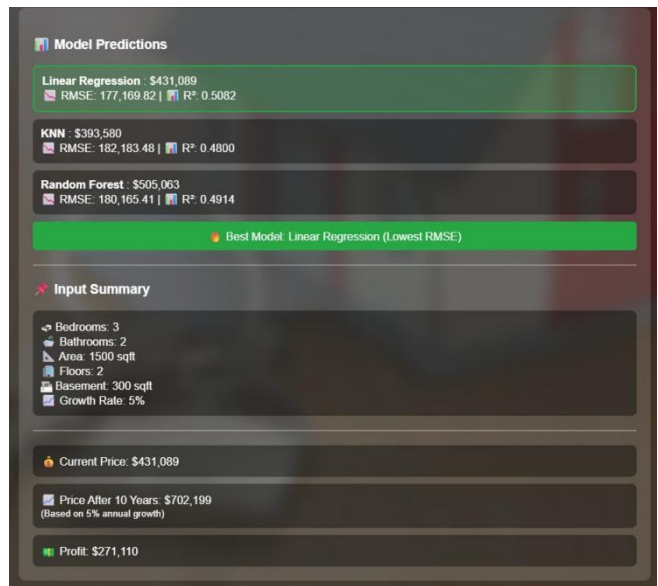


Figure 7: Predict Page (Results)

The integration of Tableau visualization and machine learning prediction provides an effective platform for analyzing housing data and estimating property prices.

## 6. Discussion

### 6.1 Contribution of Tableau Visualization

The use of Tableau dashboards enables the transformation of raw housing datasets into interactive visualizations, allowing users to analyze price distributions, property features, and overall market trends. This approach improves the understanding of complex housing data compared to static reports.

### 6.2 Machine Learning Prediction Performance

Machine learning models such as Linear Regression, Random Forest, and Gradient Boosting predict house prices with good accuracy using features like area, bedrooms, bathrooms, and renovation status.

### 6.3 Integration of Visualization and Prediction

The system combines Tableau visualization for trend analysis with machine learning models for accurate house price prediction.

### 6.4 Comparison with Traditional Housing Analysis

Unlike traditional manual analysis, the proposed system provides data-driven insights and automated price prediction using visualization and machine learning.

### 6.5 Web Integration and Deployment

The system is deployed as a web application where users can explore Tableau dashboards and input property details to obtain predicted house prices.

## 7. Limitations

Despite the effectiveness of the system, some limitations exist:

- The dataset may not include all real estate factors such as neighborhood quality and economic conditions.
- Real-time housing market data integration was not implemented.
- External economic factors like interest rates and inflation were not considered.
- Advanced deep learning models were not explored.

## 8. Future Scope

The system can be improved in several ways:

### 8.1 Deep Learning Models

Using models like ANN and LSTM can make predictions more accurate.

### 8.2 Real-Time Data

Adding real-time housing market data can give up-to-date predictions.

### 8.3 Location Features

Including details like neighborhood quality, nearby facilities, and infrastructure can improve accuracy.

### 8.4 Cloud Deployment

Hosting the system on the cloud can make it accessible anytime and support real-time predictions.

## 9. Conclusion

This research presents a data-driven system for housing market analysis and house price prediction by integrating Tableau visualization with machine learning techniques.

The system analyzes housing trends through interactive dashboards and predicts property prices using machine learning models. Experimental results show that features such as area, bedrooms, bathrooms, floors, and renovation status significantly influence house prices.

The integration of visualization and prediction provides an effective platform for analyzing housing markets and supporting data-driven decision making.

## 10. Acknowledgement

The authors express their sincere gratitude to M. Satish, M.Tech, Assistant Professor, Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning), Kallam Haranadha Reddy Institute of Technology, for his continuous guidance, valuable suggestions, and encouragement throughout the development of this project.

The authors also thank the Head of the Department, faculty members, and laboratory staff of the Department of CSE (AI & ML) for providing the necessary support and academic environment to successfully complete this research work.

Special appreciation is extended to all team members for their contributions during dataset preparation, dashboard development using Tableau, machine learning model implementation, and integration of the prediction system through the web application.

Finally, the authors acknowledge the support of open-source platforms and tools such as Python, Scikit-learn, and Tableau, which made the successful implementation of this project possible.

## 11. References

1. A. S. Ravikumar, "House Price Prediction Using Machine Learning," National College of Ireland, 2018. <https://norma.ncirl.ie/3096/1/aswinsivamravikumar.pdf>
2. B. Ibok, "Big Data Analytics and Data Visualization: Predicting Property Prices in Perth Using Big Data Analytics – A PySpark and Tableau Approach," 2024. <https://www.researchgate.net/publication/391700673>
3. A. Kudavkar, "House Price Prediction Using Machine Learning Algorithm," ResearchGate, 2021. <https://www.researchgate.net/publication/350430324>
4. S. Akhil, N. Kaur, "House Price Prediction Using Machine Learning Techniques," International Journal of Computer Science Trends and Technology, 2021. [https://d1wqtxts1xzle7.cloudfront.net/67039858/IJCST\\_V9I3P1-libre.pdf](https://d1wqtxts1xzle7.cloudfront.net/67039858/IJCST_V9I3P1-libre.pdf)
5. L. Breiman, "Random Forests," Machine Learning, 2001, 45 (1), 5–32.
6. T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning, 2nd ed. Springer, 2009.
7. J. Han, M. Kamber, J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Morgan Kaufmann, 2011.
8. C. M. Bishop, Pattern Recognition and Machine Learning. Springer, 2006.
9. I. Goodfellow, Y. Bengio, A. Courville, Deep Learning. MIT Press, 2016.
10. Scikit-learn Developers, "Scikit-learn: Machine Learning in Python," 2025. <https://scikit-learn.org>
11. Tableau Software, "Tableau Data Visualization Platform," 2025. <https://www.tableau.com>
12. Python Software Foundation, "Python Programming Language," 2025. <https://www.python.org>
13. A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow, 2nd ed. O'Reilly Media, 2019.
14. D. J. Hand, H. Mannila, P. Smyth, Principles of Data Mining. MIT Press, 2001.