



vkeya Update README.md

now



80 lines (65 loc) · 5.5 KB

Preview

Code

Blame

Raw



Business Understanding

The real estate agency aims to provide accurate property valuations to homeowners looking to sell their homes and to buyers interested in purchasing properties. Property valuation is a critical aspect of real estate transactions as it directly impacts pricing decisions, marketing strategies, and negotiation outcomes.

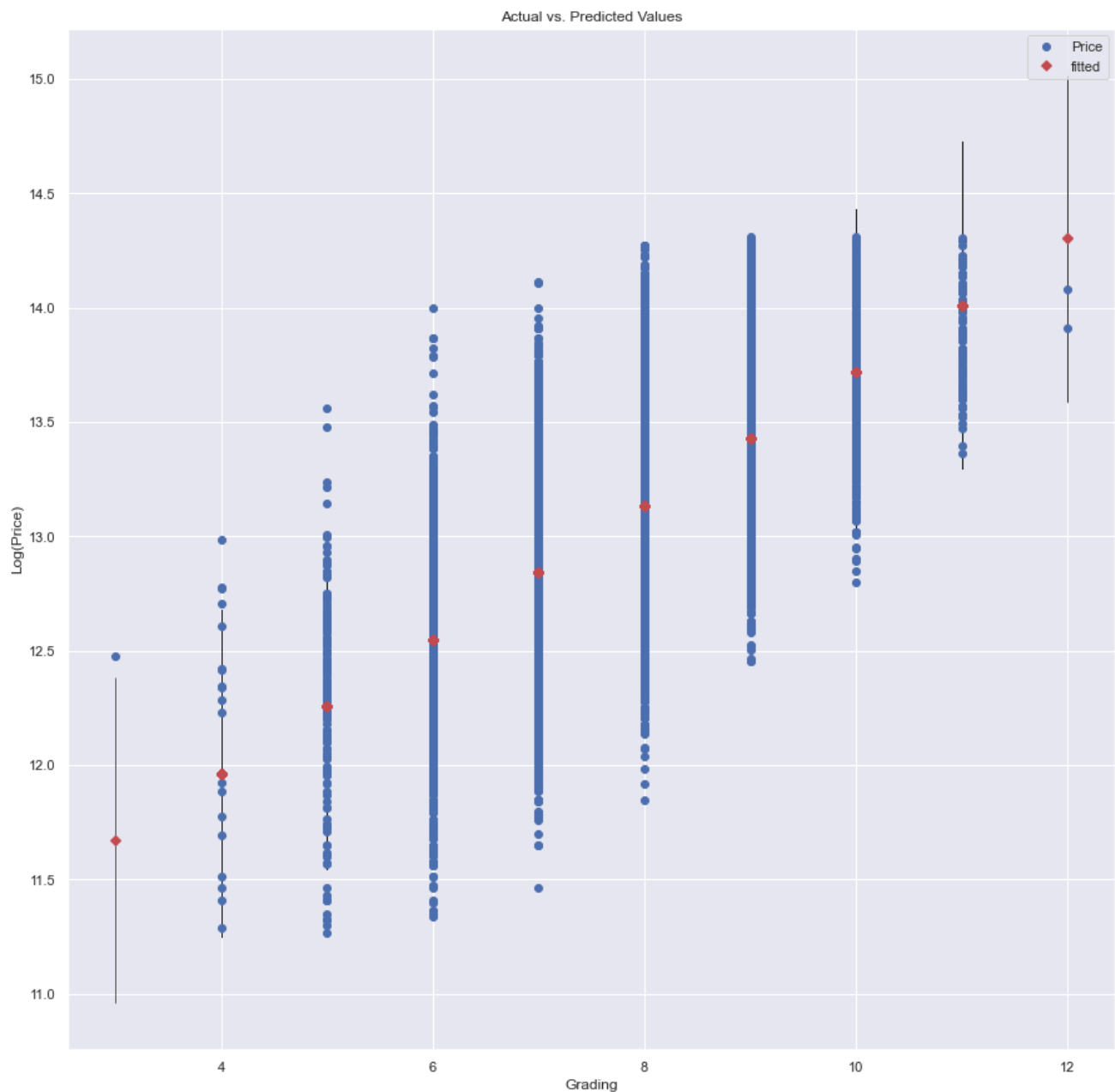
Problem Statement

The real estate agency faces challenges in accurately valuing properties, which can lead to overpricing or underpricing homes, affecting client satisfaction, time on market, and overall business performance.

OBJECTIVES:

1. Determine the key factors such as square foot living, the number of bedrooms and bathrooms, the condition of the house and others that significantly influence the house prices.
2. Develop a model that can accurately predict house prices based on these factors.
3. To determine which seasons have the highest sales.
4. To provide valuable insights for real estate agents, property developers, and investors in the company's portfolio to make informed decisions regarding pricing, renovations, and marketing strategies

Baseline model interpretation

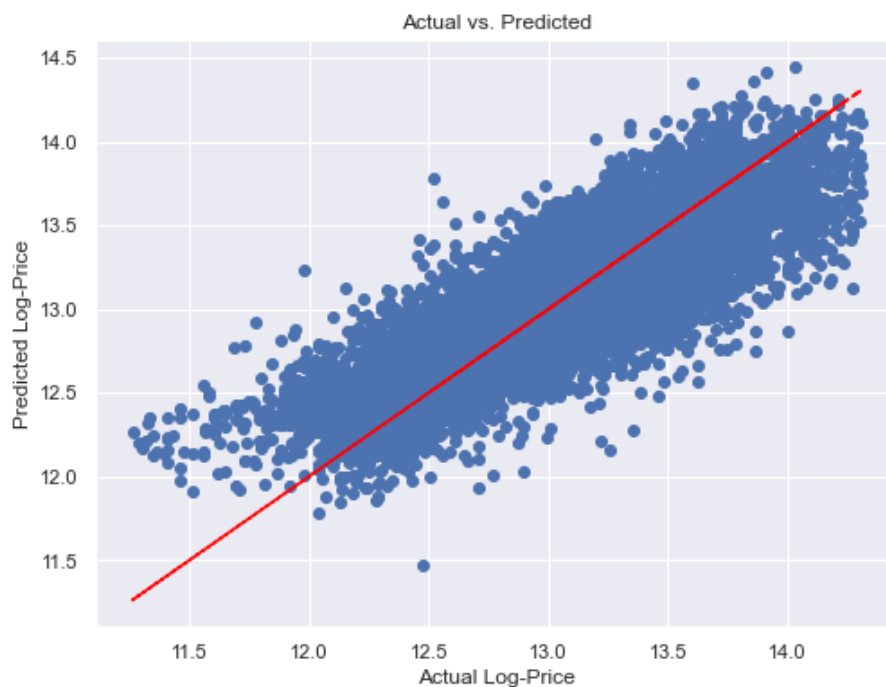


Our linear regression model is $\text{Price} = 0.2923 * \text{Grading} + 10.7957$. With a significance level of 0.05, our linear regression model is significant with a p value F statistic of 0.0. The model explains about 42% of Price as indicated by the R squared. Both our intercept and our coefficient for Grading are statistically significant. Our intercept is about 10.7957, meaning that a home with 0 Grading area would cost about 10.7957. Our coefficient for Grading is about 10.8, which means that for each additional Grading, we expect the price to increase about \$0.2923.

Evaluating Model Performance.

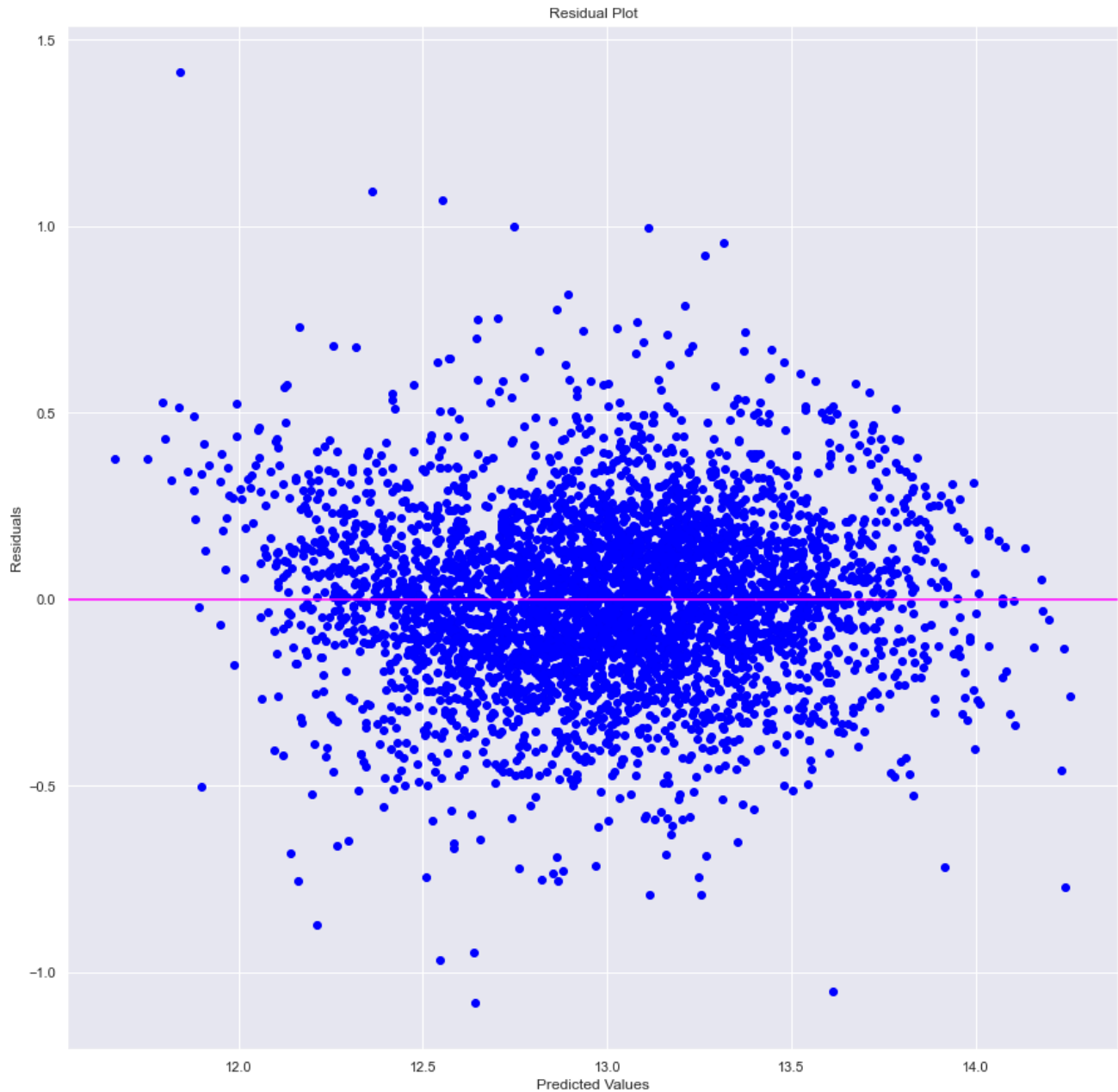
We used f statistic p value to test for significance. We also used R squared and RMSE.

Second Model Interpretation(Multiple_Linear_Regression)



The model explains 69% of the variance in price as indicated by R squared. Therefore, the multiple linear regression is better than our baseline regression.

polynomial regression



Given that polynomial regression accounts for roughly 76% of the volatility in house prices—a percentage somewhat greater than that explained by multiple linear regression—it is clear that polynomial regression is a superior model in this particular situation. Polynomial regression also has a less Root Squared Mean Error than Multiple Linear Regression Model. This suggests that the polynomial regression yielded superior results. The RMSE of the polynomial regression also suggests the same since it is also lower than the multiple linear regression model.

Conclusions

In this project, we conducted an in-depth analysis of the King County (KC) house dataset to predict house prices using advanced regression techniques. The goal was to provide valuable insights for stakeholders in the real estate industry. Key Findings: • Feature Importance: Through regression analysis, we identified several key predictors that significantly influence house prices, including grading, square footage of living space, number of bedrooms, and bathrooms. Effect of Features: • Effect of Features: Our findings reveal nuanced relationships between house features and pricing. For instance, higher grading and larger living spaces positively impact prices. • Model Performance: We compared different regression models and found that polynomial regression outperformed multiple linear regression model, capturing non-linear relationships more effectively and yielding higher predictive accuracy. on price followed by selling in winter, summer, fall and grade of house

Limitations

Data Timeliness: The dataset may not reflect the current market conditions or housing trends due to its age. Real estate markets can exhibit temporal dynamics that require up-to-date data for accurate predictions. Local Market Dynamics: The dataset's focus on King County may limit the model's generalizability to other regions or markets with different characteristics. Interpretability of Results: Some advanced modeling techniques (e.g., polynomial regression) may produce complex results that are difficult to interpret or explain to stakeholders

Recommendations

The square footage of living space has a significant influence on house pricing as well. This information can be used by the Real Estate Agency to support higher listing prices for homes with larger square footage. There are several benefits to using a polynomial regression model instead of a typical linear model when predicting property prices. By taking into consideration the non-linear effects of important features like square footage, location, and on-site facilities, this method enables us to capture intricate correlations and fluctuations in home pricing. The polynomial regression's high R-squared value indicates that the chosen features account for a significant amount of the variability in home prices, which translates into more accurate and consistent price projections. Using this cutting-edge modeling technique gives us, as real estate market participants, more decision-making power and improves pricing tactics, investment analyses, and market competitiveness overall.