PROJECT-4

Floors & fair pricing

**Summary:** My Project investigates how different floor covering types influence the price per square foot for residential properties using regression-based approach. Firstly, I cleaned the data and handled missing values and outliers. taking about Floor covering feature it is a categorical type having single and multiple type of floorcovering used for one house.so I applied one-hot encoding to handle multiple flooring types per house. Then I built a Multivariate Linear Regression model to avoid noise good performance of the model. I used different continuous features like lot size, taxes many more to predict price\_per\_sqrt. Though the model yielded a low R\*\*2 score of 0.01, it revealed useful insights for both buyers and sellers and highlighted opportunities for future work using more advanced models and features.

**Key Libraries/Techniques Used & Outcomes:**

1. **Pandas** – For data manipulation, handling missing values, and splitting multiple floor types into lists.
2. **NumPy** – Used for mathematical operations
3. **One-Hot Encoding** – Converted complex floor\_covering combinations into usable binary columns.
4. **Data Cleaning Techniques** – Removed extra spaces, standardized text, and filled missing values to create a clean dataset.
5. **Feature Engineering** – Created a custom price\_per\_sqft metric to normalize home prices for better comparison.
6. **Model Evaluation (R² Score)** – The regression model had limited predictive power (R² = 0.01), but provided valuable insights and a foundation for future enhancements.