**Housing Prices Prediction with Python**

This project applies practical exploratory data analysis (EDA) techniques on a housing

prices dataset using Python's libraries like Seaborn and Matplotlib to visualize data.

Machine learning techniques are employed to predict housing prices.

**Insights**

**Missing Data**: There are missing values in columns such as Alley and FireplaceQu,

which need to be handled appropriately during preprocessing.

**Street Types**: Most properties are on paved streets (Pave),

while a minority are on gravel streets (Grvl).

**Property Type**: Single-family Detached homes show the highest sales volumes.

**Zoning**: Residential Low Density zones demonstrate the best sales figures.

**Street Type Impact**: Paved streets have a more significant impact on sales compared to gravel streets.

**Property Characteristics**: Features like property shape and land contour significantly influence sale prices.

**Age Relationship**: There is a negative correlation between property age and sale price,

indicating that newer properties tend to sell for higher prices.

**Living Area**: There is a positive relationship between living area and sale price

, suggesting that larger homes tend to be more expensive.

**Conclusion**

This EDA highlights various aspects of the dataset, including the impact of street type,

property characteristics, and living area on housing prices.

Visualizations using Seaborn and Matplotlib help uncover relationships and distributions within the data,

providing valuable insights for further analysis and decision-making.

**Data Preprocessing**

Several preprocessing steps are performed to clean and prepare the data for modeling:

Handling Missing Values: Used SimpleImputer to fill missing values in columns.

Feature Scaling: Applied StandardScaler to standardize numerical features.

Categorical Encoding: Used OneHotEncoder to encode categorical features.

Column Transformation: Combined preprocessing steps using ColumnTransformer.