Residential Real Estate Price Prediction ~ Preliminary Summary

Real Estate Investment Trust • Investment Analytics Project

OVERVIEW

- The Real Estate Investment Trust is exploring opportunities in the residential housing market. As part of this initiative, an initial review of historical house sales data was conducted to ensure the data is reliable and suitable for further analysis.
- This preliminary step focused on understanding the structure, completeness, and general characteristics of the data before moving into deeper analysis and modeling.

PROJECT STATUS

- The housing sales data has been successfully loaded into the database and reviewed for accuracy and completeness.
- Initial checks confirm that the data is well-organized, consistent, and appropriate for understanding residential property prices.
- At this stage, the dataset has been validated and is ready for deeper exploration in the next phase of the project.

NEXT STEPS

- Conduct a full **exploratory data analysis (EDA)** using Python.
- Examine variable distributions and identify potential outliers.
- Explore relationships between key predictors and house prices.
- Prepare features for regression-based and machine learning models.

KEY INSIGHTS

- No missing values were detected across any columns, indicating high data completeness.
- Price values range from approximately **$78,000 to $7.7M**, which is realistic for residential properties.
- Living area sizes vary from **~370 to ~13,500 sqft**, covering small to luxury homes.
- Average house prices increase consistently with:
  - Number of bedrooms
  - Larger living area categories
- Waterfront properties represent a small subset of the data but command a **significant price premium** compared to non-waterfront homes.
- Structural variables such as grade, square footage, and location coordinates appear well-suited for predictive modeling.
- Overall, the dataset contains strong candidate features for estimating house prices.