Project Background

Assuming you are acting as a consultant for a Real Estate Investment Trust (REIT).  
The REIT invest in houses and apartments in New York state

Part of the REIT business is try to predict the fair transaction price of a property before it’s sold   
They invest in houses, apartments, and condos within a small county in New York state.  
They do so to calibrate their internal pricing models and keep a pulse on the market.

Current Working Solution for the Tasks

The REIT currently employs a third-party appraisal service for estimating the price of the property with their own expertise.

In practice, the skill level of individual appraisers vary quite large.

To estimate the mis-priced range, the REIT run a trial run to compare the actual transaction prices to the estimates from the appraiser.

It was found that the estimates given by inexperienced appraisers differs **$70,000** on average

Our role as a consulting data scientist

The REIT has hired us to find a data-driven approach to valuing properties instead of relying on the personal expertise from the appraiser  
The REIT currently have an untapped dataset of transaction prices for previous properties on the market. Our task is to build a real-estate pricing model using that dataset.  
If we can build a model to predict transaction prices with an average error of under $70,000, then our client can replace inexperienced appraisers with our model.

Problem Specification and Scope

**Deliverable**: Trained model file  
**Machine learning task:** Regression  
**Target variable**: Transaction Price  
**Win condition**: Avg. prediction error < $70,000, using Mean Absolute Error (MAE)