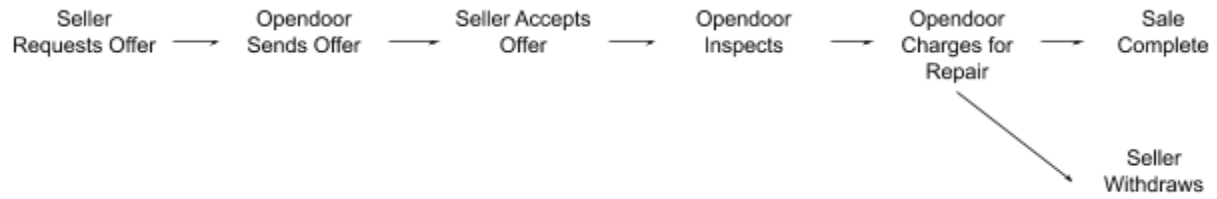
Getting Upfront Repair Costs

Prompt:
Inspections and repair estimation are a part of the current home seller process (see diagram below). Sellers may request an offer from Opendoor. Upon offer acceptance, an Opendoor estimator visits the home to scope common repairs needed before listing.

The common scopes are then charged to the seller as a requirement for the home sale. For example, if the estimator identifies $1500 in common scopes, the seller can either decide to accept the $1500 charge and complete the sale or withdraw from the process. For every $150 in repair charge we can expect to have a 1% withdrawal rate.



We want to manage our withdrawal rates because every home Opendoor acquires translates to $5000 in profit on average. One option is to eliminate inspections and tell sellers how much repairs will likely cost as soon as they request an offer. We think this added convenience may reduce withdrawal rates by 5 percentage points.

Please analyze the attached data to determine whether we should provide an estimate of repairs at time of offer. If we should not, please explain why. If we should, please propose a roadmap to do so with the following milestones:

- V0 solution (1 Month)
- V1 solution (1Quarter)

Deliverables:

1. **Notebook** showing analysis and results

2. **Slide deck** presenting results and recommendations for technical and non-technical audiences (Product manager, operations leaders, and data scientist peers). If you progress to onsite you will be asked to present your findings and may use this slide deck.

3. **csv with your model's predictions** on the holdout set of homes assessed in the future. We will use your predictions to assess the accuracy of your final model.

Please submit a csv (commas delimit columns, newlines delimit rows) with a prediction for each row in the holdout set. The first row should be a header, the first column should be the `flip_token` of each home, and the second column should be the `prediction` in USD. The total length of your submission should be 1580 rows. Please see the included `example_submission.csv` sample submission. Link: example_submission.csv

Grading Areas:

- Software Engineering/Coding Practices
- Statistical Analysis
    - Data Exploration
    - Analytic Approach
    - Modeling Strategy
- Business Judgement and Leadership Qualities
    - Potential for Impact
    - Complete and cross-functional solution
    - Incorporation of analytic findings into business presentation

Data Notes:

The data for this exercise comes from a sample of homes inspected by Opendoor. The column **common_scope** is the scope of common repairs referenced in the prompt above. All other data can be assumed to be collected at the time the seller requests an offer:
There are 2 datasets:

1. **development_df** - the labeled dataset with historic data
2. **holdout_candidate_df** - the unlabeled dataset simulating homes we will buy in the future. If you develop a model, use this dataset to generate and submit predictions for model evaluation.