**Problem Description**

We work for a national real estate investment company that specializes in "flipping" residential and non-residential property. In the past we have made our investment decisions based on the intuition of our researchers. To expand our operations throughout the country, we often purchase property remotely, without being able to send an agent and assess its true value. Additionally, the budget for repairs and general improvement is based on recommendations from the local contractors we hire in the area of the newly purchased property.

We have kept data on our previous investment decisions. In order to scale our process and reduce the bottleneck of our researchers, we would like you to build us a model/models or a rule set that allows us to maximize our profit. There are two ways in which our business makes profit. One is by purchasing properties that are undervalued (this is when initial\_price is less than initial\_value). The other is by investing money into improving and upgrading the properties we have purchased and having the market value (or final price) increase more than our investment. Your approach will be used to decide which properties to purchase and how much should be invested in their improvement.

When you have a model built, you will use it to decide which of the properties in the validation data we should purchase and how much money should be spent upgrading them. Your budget is $400,000,000 for purchasing the properties and $100,000,000 for upgrading them. These budgets are not interchangeable as they come from different sources.

For each property\_id, fill out the purchase\_decision column with a 1 if you recommend buying it and a 0 if you do not. For the properties you recommend purchasing, also fill out the investment column with the amount of money that we should spend upgrading the property. Remember, the sum of initial\_price for purchased properties cannot exceed your $500,000,000 budget; as well, the additional investments made improving them cannot exceed your $100,000,000 budget. We will process the rows in the order provided and stop once you have spent your budget. Profit will be calculated once, so you cannot use profit from one sale to fund the purchase of another property.

Given your level of additional investment, we will calculate the final sale price (final\_price) for the properties you choose to purchase. Your overall score will be the total amount of profit that your decisions generate. Profit is defined as the sum of [final\_price - investment - initial\_price] for all of the properties purchased. The better you are at identifying under priced property and the more efficiently you set a budget for our contractors, the higher the return will be on our investment.

A helpful note: There is likely not enough time to solve this problem perfectly. I would recommend thinking about how to get to a solution somewhat quickly and then iterate improvements to your approach.

**Column List**

**Generel Information**

property\_id: unique identifier

zone: general property type (commercial, residential, industrial, and mixed-use)

sub\_type: more specific classification of property (restaurant, single family home, vacant lot, etc.)

**Property Location**

street\_name"

street\_number"

address\_line\_2"

city\_name"

zip\_code"

**Relevant Dates**

days\_on\_market: number of days the property has been listed as for sale

build\_date: the date the property/lot was originally built or parceled

remodel\_date: the date the property was last remodeled or improved

**Neighborhood Information**

area\_type: type of surrounding area (suburban, rural, or urban)

current\_population" .= current population of town/city

population\_5\_years\_ago: population of town/city 5 years ago

schools\_in\_area: number of schools in the town/city

public\_transit\_score: measure of the availability of public transportation in the surrounding area (higher score corresponds to higher transportation availability)

crime\_score: measure of the amount of crime in the surrounding area (higher score corresponds to higher crime rates)

culture\_score: measure of the quality of culture/lifestyle in the surrounding area (higher score corresponds to higher cultural quality)

average\_neighborhood\_price: average price of similar properties in the area

**Condition**

damage\_code: a designation of any serious damage affecting the property (fire damage, flood, mold, etc.)

inspection\_type: the type of inspection that was conducted on the property (inspected by seller, buyer, foreclosure agent, etc)

structural\_quality\_grade: the inspector's score for the condition of the structure/land

exterior\_condition\_grade: the inspector's score for the condition of the property's exterior

interior\_condition\_grade: the inspector's score for the condition of the property's interior

utilities\_grade: the inspector's score for the condition of the utilities on/in the property (electric, water, sewage, HVAC, etc)

damage\_and\_issue\_grade: the inspector's score for any damage or miscellaneous issues found with the property

overall\_inspector\_score: a overall score between 0 and 100 assigned by the inspector

**Qualities**

sqft: square footage of structures or acreage of vacant lots

floors\_in\_building: the total number of floors in the building (may not all be part of the sale in the case of condos, retail, etc)

floors\_in\_unit: the number of floors of the property that is for sale

floor\_of\_unit: the floor of the building that the property is on (if the property is a part of a larger, multi unit structure)

bedrooms: number of bedrooms in the property

bathrooms: number of bathrooms in the property

parking: availability of parking at the property

basement: flag representing the existence of a basement

central\_hvac: flag representing the existence of hvac

misc\_features: a designation of any important features associated with the property

exterior\_color: color of the structure

exterior\_material: building material of the structure's frame

**Pricing**

initial\_price: the asking price for the property (assumed non-negotiable)

initial\_value: the actual value of the property determined by our assessor (left blank in the test data)

**Choice Variables**

purchase\_decision: a flag representing the decision to purchase a property or not (left blank in the test data)

investment: the amount of money spent fixing damage and making general improvements to the property (left blank in the test data)

**Outcome**

final\_price: the final sale price of the house after we have invested in improvements (determined by our scoring code as a function of property value and investment amount. left blank in the test data)