Instructions

Objective

You own a diamond business and you have to replenish your inventory. You have $5,000,000 budgeted towards the purchase of new diamonds. You have a historical list of diamonds that were purchased from various distributor vendors and the prices they retailed for. You also have a list of diamonds that are currently on the market, but this list does not include any set prices. Instead, you will place an offer for a diamond and the distributor will make a single yes-or-no decision to sell you the diamond at that price. If your offer is accepted then you receive the diamond and you will be awarded the difference between the retail price your company can eventually sell the diamond for and the price that you offered as profit. If you fail to make a purchase, you simply do not get anything for profit.

What you will return to Enova is a copy of any code used, a description of your process, answers to the short answer questions, and the offers file with the offer column populated. The offer column should be populated only on diamonds you wish to make an offer on, and the sum of the offers needs to be less than or equal to $5,000,000.

Directions

You have been given two data sets. The dataset labeled "training.csv" contains historical data regarding how much diamonds have sold for; both the retail price the diamond sells at and the price paid to the vendor to acquire it. You can use this data to train any models you see fit to answer the question. The second data set labeled "offers.csv" is the dataset you will provide offers for. Note neither the retail price or the price you would need to pay is present in the offers file.

Also included is a data dictionary explaining the fields in the datasets.

Evaluations of offers

Recall that your max budget is $5,000,000 of offers, so the total sum of your offers needs to be less than $5,000,000. You will be charged the value of your offer even if the seller would have been willing to sell for less (You always pay your offer if it is accepted). If your offer is rejected for being less than the price the seller wants you pay nothing, but you don't receive the diamond.

Your profit will be calculated as the sum of retail price the diamonds will eventually sell for minus the sum of your accepted offers. You make no profit on rejected offers.

Short Answer Problems

Please provide short (a paragraph or so) answers to these questions. Feel free to include any charts or graphs that you feel help answer the questions.

1. Do you think that any of the vendors are over or under charging for diamonds compared to the other vendors? Do you think they are selling the same kind of diamonds? How did you come to this conclusion?
2. What is the relationship between Carat and Retail Price? Why do you think the relationship takes this form?

Submission

A full submission contains:

- The csv of the offers dataset with your offers. The 'id' column must be unchanged and included; the offers should be in the column 'offers'. The sum of the offers must be less than or equal to $5,000,000.
- A copy of any code used in to solve the problem.
- A brief write-up detailing what you did and why. If you feel that your code is sufficiently documented explaining your process that will satisfy this requirement. Including graphs and charts that helped you is a great way of showing your thoughts and process, but are not necessary.
- Brief (a paragraph or so) answers to each of the short answer problems.

The amount of profit you make, the methodology you use, and your answers to the short questions will be taken into account by Enova.

Scoring

We will measure the performance of your model/rules by profit. In other words, we will take predictions from your scored data and compute the following:

$$\sum (Retail\ Price\ of\ Accepted\ Offers - Accepted\ Offers)$$

This being said, profit is not the only aspect of your model we will examine. We care about your approach and knowledge more and weigh your write up and description just as heavily as your final profit.

Tips

- Be sure to examine the scoring data before beginning and reflect on what data might be important to consider when solving the problem.
- The datasets contain a mix of numbers and text. Look it over, consider how you might extract meaningful information from text values.
- Your budget limits the total amount you can offer even if your offer isn't accepted.
- Rejected offers don't return any profit, but do count towards your budget.
- Feel free to perform outside research to better understand what relationships might exist in diamond pricing
- If you don't know how to build a statistical model, that's fine! Try to think about a set of rules that can help you figure out which diamond to put offers on. A good, data-driven, rules-based approach can show your skill with data as well.
- The write-up is your chance to walk us through, not just the exact steps you took to solve the assessment, but the journey you took to get there. For example, feel free to include details on why you selected a particular method or anything you tried that didn't work.
- Previous candidates have attempted to run the data through a black box or generic model before thinking critically about the data and problem. They typically saw limited results

Thank you for your interest in Enova! We look forward to speaking with you. Good Luck!