Guided Capstone Project Report

In order to best answer the question of weather Big Mountain resort was maximizing their pricing and provide assistance in park investment decision making, a random forest regression model was developed to predict how a ski resort with certain attributes could charge. This model was created after data wrangling (removing rows of data with missing or questionable values/columns of data that provided no useful information), exploratory data analysis (where PCA was performed to determine the best variables to use in our model), and preprocessing and training (where we decided on using a random forest regression model due to its low cross-validation mean absolute error and lower variability compared to other options).

Using this model, it was found that while Big Mountain currently charges $81.00, a charge of around $95.00 could be supported by the current marketplace. Big Mountain has a lot of key features that are better than most other resorts. Installing the additional chair lift would lead to estimated revenue increase of about $350k per year. Since revenues increases were marginal with increasing snow area and increasing largest run, I would avoid changing those features. Overall I would recommend adding the chair lift to make for a longer run (due to the expected revenue increase) and offsetting that expense by closing between 1-5 other runs, due to the estimated revenue loss of 0 for closing 1 run and just over 100k for closing 5 runs (see Figure 1).

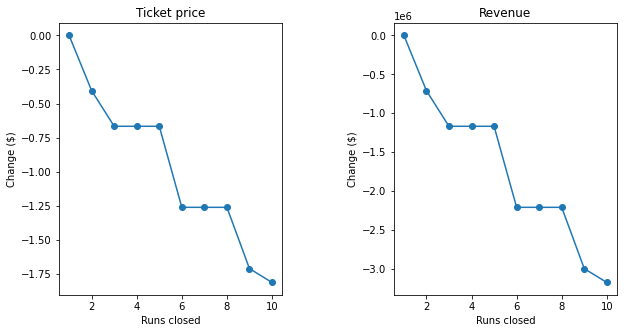


Figure 1. Ticket price and Revenue predictions (in the 100,000s) as Run closures occur.

I recommend looking into how much maintenance cost there is for keeping each run open in order to maximize the run maintenance/revenue tradeoff. The data used for this model was not perfect, there were a good amount of missing values, and there was only price data on ticket prices. Price data on other features such as rentals, food, lodging, etc. would be very helpful as that would likely significantly affect how much an individual costumer affects revenue at the resort overall. It is possible that because Big Mountain has features that when combined make it a top resort, that its relatively high prices are still too low. While this may come as a surprise for the business leaders, it should be emphasized that a model should be taken as an educated estimate, and could very likely not support the reality of ticket prices. This could be tested only by adjusting ticket prices and finding out how it effects revenue. Assuming this model was useful, I could make an interface for the model where non data analysts could simply enter various values into labeled tabs and a function would give the predicted revenues.