Introduction

The goal of this project is to assist Big Mountain Resort in the re-evaluation of their ticket pricing strategy in order to increase revenue next year. Their current ticket price of $81 is based on charging a premium over the average ticket prices of resorts in its market segment. Pricing based on market average obscures which ski resort features are important to visitors and makes it difficult to develop an investment strategy. There is already a sense that Big Mountain is underutilizing its existing facilities, so it is essential to understand which features are most important to resort visitors in order to maximize utilization and determine if there is justification for a ticket price increase. We have used data science and machine learning to develop a model that enabled us to better understand the ski resort market in which people pay more for certain facilities, and less for others, and we were able to predict which facilities at Big Mountain are able to support an increase ticket price.

Business Opportunity

Big Mountain Resort has been reviewing potential scenarios for either cutting costs (without undermining the current ticket price) or make changes to the resort that would support a higher ticket price. We have used our model to predict the impact that these changes would have on ticket price and revenue. The four scenarios currently under consideration are as follows:

1. Permanently closing down up to 10 of the least used runs.
2. Increase the vertical drop by adding a new run to a point 150 feet lower. This would require the installation of an additional chair lift to bring skiers back up.
3. Same as scenario 2, but with 2 additional acres of snowmaking cover.
4. Increase the longest run, Hellfire, by 0.2 miles for a total length of 3.5 miles. This would require additional snowmaking coverage of 4 acres.

Analytical Framework

Data

The data set received from the Big Mountain database manager contained data for 330 ski resorts in the US, including Big Mountain. The resorts are located across 35 states, comprising 38 unique regions. The original dataset contained the adult weekday and weekend tickets prices of the resorts, as well as 22 other quantitative variables of resort features of that a visitor might find of value, such as the count of the different types of and passenger capacity of transport across the resort, skiable terrain, snowmaking coverage, number of runs, and vertical drop. 14% of the resorts did not have data for adult weekend or weekday ticket prices and were removed from the dataset since ticket price is our target feature. This left 277 resorts for further analysis. In order to determine if there were any state-specific effects that affect ticket pricing, population and area data for the US states was gathered from Wikipedia, transformed, and added to the data set in the form of the number of resorts per 100k capita and per 100k square miles, respectively.

Analysis

Because we have so many feature variables, including the non-quantitative state value, I used a method called principle components analysis to better understand a) the source of the variation between states, and b) the dependence or independence of these variables. PCA produced no evidence of a clear pattern among the states, so the pricing model that was constructed that considers all states together.

For the machine learning phase of the analysis, I removed Big Mountain from the data set and split 70% of the dataset into a training set for modeling and the remaining 30% were reserved as a test set for testing the models. I used two types of algorithms come up with models to determine which features mostly strong influence ticket pricing and that can be used to evaluate Big Mountain's proposed changes- simple linear regression with cross-validation and the random forest regressor. According to linear regression analysis, the 8 best features, in order of feature strength, are as follows: 1) length of vertical drop, 2) snowmaking acreage, 3) the total number of chairs, 4) the number of fast quads, 5) the number of runs, 6) the length of the longest run, 7) the number of trams (negative correlation), and 8) the acreage of skiable terrain (negative correlation). The strongest top four features predicted by random forest, in order of strength, are as follows: 1) number of fast quads, 2) number of runs, 3) snowmaking acreage, and 4) length of vertical drop. In order to determine which of the two models to select, I calculated the mean absolute error of the models, which measures the average magnitude of the errors in a set of predictions. The random forest model has a lower CV mean absolute error by almost $1 and exhibited less variability overall. For this reason, I opted to move forward with the random forest model.

Results

We used our model to predict Big Mountain's expected ticket price based on its current features. The model price for the resort is $95.87, with a MAE of $10.39. This means we might expect to be off by around $10.39 (between $85.48 and $106.26) if we guessed ticket price based on an average of known values. This suggests that even without making any changes to the facilities, a ticket price increase would be tolerated by the market.

We used the model to further predict the effect that the four proposed scenarios would have on ticket price and revenue. In order to do so, we used historic data to estimate the number of expected visitors per season (350,000 adults) and the average number a days skied per visitor (5 days).

1. Close up to 10 of the least used runs. The model suggests that:
    - Closing 1 run will have no impact on ticket price or revenue.
    - Closing 2 runs reduces support for the current ticket ($0.40 less) and therefore, revenue would likely decrease
    - Closing 3-5 runs further reduces support for the current ticket ($0.67 less) and therefore, revenue likely would decrease further. (If 3 runs are closed, Big Mountain may as well close down 4 or 5 as there's no further loss in ticket price/revenue)
    - Closing 6 or more runs ($1.26 to $1.81 less) will lead to a large drop in recommended ticket price, and therefore, revenue.
2. Increase the vertical drop by 150 feet and install an additional chair lift. No additional snow making coverage. The model provides support for:
    - A ticket price increase of $1.99
    - Over the season, this could be expected to amount to a revenue increase of $3,474,638.00
3. Increase the vertical drop by 150 feet, install an additional chair lift, and add 2 acres of snowmaking coverage. The model provides support for:
    - A ticket price increase of \$1.99
    - Over the season, this could be expected to amount to a revenue increase of \$3,474,638.00
4. Increase the longest run by 0.2 mile and add 4 acres of snow making coverage. The model does not provide any support for an increase in ticket price, and therefore no revenue increase would be expected.

Conclusion

Recommendation

Big Mountain currently charges $81 dollars for an adult weekday or weekend ticket. Our modelling suggests that you charge an even higher ticket price and not expect to experience any decline in market value. The model also provides support for a $1.99 price increase in the event that you choose to increase the vertical drop by adding a run to a point 150 feet lower down. This would require the installation of an additional chair lift, thereby increasing Big Mountain's operating expenses by $1.54 million next season. If Big Mountain is able to sustain historic levels of visitation next season, you would expect a revenue increase of $3,474,638.00 with the higher ticket price, which could be used to offset the increased operating expenses of a new chair lift. Moving forward, I would recommend increasing the vertical drop with an additional chair lift, but without increasing snowmaking coverage.

Next Steps

It might be worth investigating further which of the least-used runs could be closed down, as this would not be expected to impact revenue. There could be hidden cost-savings in the closure and/or resources for this run could be diverted elsewhere. To that end, it would be helpful to get data usage data about the 105 runs at Big Mountain to determine which features of the runs the visitors find most valuable. It would also be advantageous to understand the spread of operating expenses for each of the runs in order to determine if those can be factored in our model as well. Big Mountain resort is at the high end of the distribution of resorts for the number of fast quads, number of runs, and snowmaking acreage, which are also the dominant features predicted by our random forest model (#1, #2, and #3 respectively). While our modeling suggests that there is an inverse correlation between total skiable acres and ticket price, I think that there may real opportunity to expand into these skiable acres by providing more guaranteed snow, new runs to ski, and a faster way to get skiers to these skiable areas. It would be helpful to get some information the cost of installation of new fast quads and runs, as well as the operating expenses associated with fast quads, runs, and snowmaking machines.