Big Mountain Resort, a ski resort located in Montana, has recently installed an additional chair lift to help increase the distribution of visitors across the mountain. This additional chair increases their operating costs by $1,540,000 this season. Here we want the best solution to balance the costs or find the most profitable way that doesn’t affect the business.

Here we have constructed some potentially useful and business relevant features, derived from summary statistics for each of the states. There some states are higher in some but not in others. Some features are more correlated with one another than others. Like New York State boasts an especially large night skiing area, and it had the most resorts but wasn't in the top five largest states. Correlation between the features leads us to decide regarding which features to use in subsequent modeling. There is some positive correlation between the ratio of night skiing area with the number of resorts per capita. In other words, it seems that when resorts are more densely located with population, more night skiing is provided. There's a strong positive correlation with vertical\_drop, fastQuads, Runs and total\_chairs with ticket price.

In the scatterplots we see what some of the high correlations were clearly picking up on. It seems that the more chairs a resort has to move people around, relative to the number of runs, ticket price rapidly plummets and stays low. What we may be seeing here is an exclusive vs. mass market resort effect; if we don't have so many chairs, we can charge more for your tickets, although with fewer chairs we are inevitably going to be able to serve fewer visitors. The price per visitor is high but number of visitors may be low. It also appears that having no fast quads may limit the ticket price, but if resort covers a wide area then getting a small number of fast quads may be beneficial to ticket price.

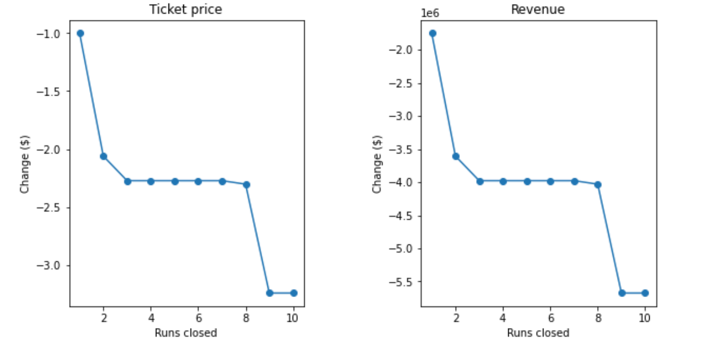
Simply taking the average price would give us the best guess or the closest answer to the predicted ticket price. There are many ways of assessing how good one set of values agrees with another, which brings us to the subjects of metrics. Here we see that our simple linear regression model explains over 80% of the variance on the train set and over 70% on the test set. Clearly, we are onto something, although the much lower value for the test set suggests we are overfitting somewhat. Using this model, then, on average you'd expect to estimate a ticket price within 9 or so of the real price. This is much, much better than the 19 from just guessing using the average.

These results highlight that assessing model performance in inherently open to variability. We'll get different results depending on the quirks of which points are in which fold. An advantage of this is that we can also obtain an estimate of the variability, or uncertainty, in our performance estimate. The random forest model has a lower cross-validation mean absolute error by almost $1. It also exhibits less variability. Verifying performance on the test set produces performance consistent with the cross-validation results.

The Big Mountain Resort has been reviewing potential scenarios for either cutting costs or increasing revenue (from ticket prices). Ticket price is not determined by any set of parameters; the resort is free to set whatever price it likes. However, the resort operates within a market where people pay more for certain facilities, and less for others. Being able to sense how facilities support a given ticket price is valuable business intelligence. This is where the utility of our model comes in.

Business has shortlisted some options:

1. Permanently closing down up to 10 of the least used runs. This doesn't impact any other resort statistics.



The model says closing one run makes no difference. Closing 2 and 3 successively reduces support for ticket price and so revenue. If Big Mountain closes down 3 runs, it seems they may as well close down 4 or 5 as there's no further loss in ticket price. Increasing the closures down to 6 or more leads to a large drop.

1. Increase the vertical drop by adding a run to a point 150 feet lower down but requiring the installation of an additional chair lift to bring skiers back up, without additional snow making coverage

This scenario increases support for ticket price by $9.45

Over the season, this could be expected to amount to $16,545,455

1. Same as number 2, but adding 2 acres of snow making cover

This scenario increases support for ticket price by $9.88

Over the season, this could be expected to amount to $17,287,879

1. Increase the longest run by 0.2 mile to boast 3.5 miles length, requiring an additional snow making coverage of 4 acres

There is no difference.

In conclusion, I would choose scenario 2 between secranios as my final option. It increases support for ticket price by $9.45 and could be expected to amount to $16,545,455.