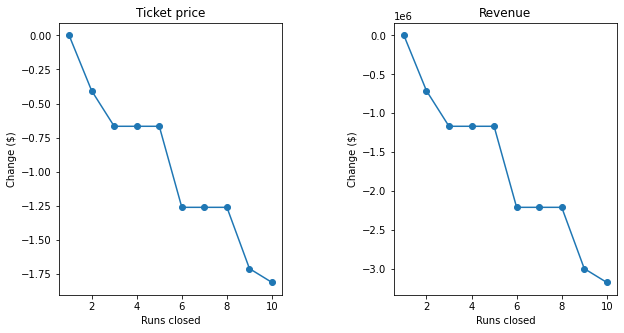
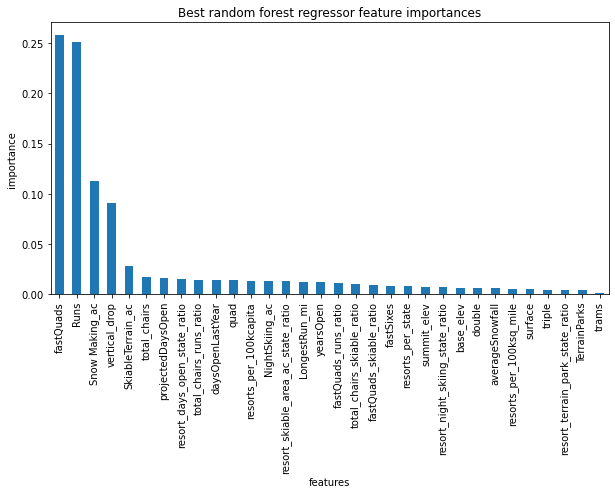
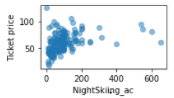
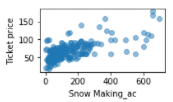
In this report, I will detail the results of the investigation of my data science team into the pricing strategy of Big Mountain Resort, why the data driven model we’ve created is worthy of trust, and how the model can be used and improved in the future.

Big Mountain’s current ticket price is $81, which is based on a flat premium charge above the average resort price in its market segment. The price does not account for the relative value of Big Mountain’s specific facilities. A random forest machine learning model created from the facility data of 276 USA resorts suggests a Big Mountain ticket price of $95.87 with a margin for error of $10.39. That’s an average increase of about $15 (another ~26 million dollars of revenue for the season), and an absolute minimum increase of about $4.50 (another ~8 million dollars of revenue for the season). Since Big Mountain is already the largest resort in Montana, I think erring towards the lower range of the suggested price increase (e.g. $5-$15 ticket price increase) is wiser since it already nets enormous profits that more than cover the costs of the additional chair lift ($1,540,000 per season) and will naturally have a more predictable result.

The model also recommends removing one run from Big Mountain. That is because having one fewer run doesn’t affect the ticket price that can be supported (meaning it adds no profit), and it costs money to maintain. It is also possibly beneficial to remove more runs, but our model can’t answer this because we weren’t given the costs of a run. If management wants to remove more runs, it’s best to remove either 5 or 8, since those are the highest numbers of runs that can be removed (saving the most money) before reaching another cutoff point that suggests a drop in ticket price, as supported by the above graph on the left.

Our model also supports the decision under consideration to add another run which will increase the Big Mountain vertical drop by 150 feet, in addition to the decision to add another chair lift to support it. Such an increase in vertical drop supports another $2 added to the ticket price, increasing revenue by about $3,500,000 for the season. This more than absorbs the cost of the added chair lift, and the management will have to decide whether it absorbs the cost of the run itself. However, our model found vertical drop to be one of the most important features for selling a ski resort ticket, so it’s very beneficial for a resort to improve it. There were also considerations to add 2 acres of snow making area to the above decision, and this is strongly recommended against by the model because the 2 acres added would cost money and not support a higher ticket price. Likewise, increasing the longest run by .2 miles with another 4 acres of snow making area is strongly recommended against, since it similarly can’t support a higher ticket price to make up for the cost of these relatively small changes.

Before training this model, we did a lot of preliminary analysis of the ski resort data and we found that features with a high correlation to ticket price are: snow making area covered, night skiing area covered, number of fast quads, total runs, total chairs, and highest vertical drop. The best features spontaneously selected by our model are number of fast quads, total runs, snow making area covered and vertical drop, with other features being relatively unimportant. Given that there were 32 total features that could have been selected, the fact that the model is heavily focused on features we identified as important is a great sign that the model is on the right track. In fact, it is good that total chairs is not included among the most important features since it is highly correlated to total runs, which would make it redundant (night skiing area covered correlates similarly in its distribution to snow making area covered, but snow making area has a clearer positive trend as resort price increases).

In addition to having good features, the model performs much better than a simple trick like taking the mean: during cross-validation, the average r-squared score it received is .71 (whereas the mean would have a score of 0), and the mean absolute error is $10.39 as compared to the $19.14 mean absolute error of the mean. This implies that the model was worth the effort. In addition, we found that there were only very small gains in model accuracy upon adding more data to training once it reached a threshold of 40-60 resort examples, so the 276 resorts we had access to was more than enough data to train the model.

Big Mountain can use the current model for any scenario analysis they are considering in the future; it should be made available to management in a production environment so that feature changes can be input to understand what ticket price those features support (which the model will output based on the suggested changes). For model improvements, more data about the cost of runs would be useful for cost-benefit analysis, in addition to data about the typical length of stays at each resort (which is likely proprietary, thus difficult to obtain, but would allow us to model features against revenue rather than just price).