Big Mountain Resort needs to optimize their pricing based on their facilities to get the most people in that will maximize revenue. They are currently basing pricing off of ‘similar’ resorts, which may not be defined properly. In determining which amenities are important to customers for the price they pay, they may also be able to strip certain amenities without impacting the value they provide to cut costs. Conversely, they are open to spending more on amenities if they justify higher entrance fees. We are attempting to determine what pricing should be for the best proposed mix of amenities and facilities.

Big Mountain Resort is losing out on potential revenue by not charging the right price for what they currently have. They may also not have the right mix of facilities and amenities in service that are important to customers when they decide to pay the price of a ticket. By using the existing data about where Big Mountain fits into the other resorts, we can figure out the best way to position Big Mountain

We began by loading up the data that was provided to us and doing a basic quality check – throwing out one-off observations that we couldn’t explain or provide better values for, making sure our observations were unique, and bringing in some other, readily available data about the states in which the resorts we are studying sit.

Once we did that, we took a look at how the attributes of each observation (resort) may fit in, with a specific bent toward state data (location seems very likely to inform population, economics, skiing type, etc.). Then we did primary component analysis to look at how much of each attribute goes into explaining how correlated the attributes are to each other.

By then building a seaborn heatmap, we were able to see Turning your attention to your target feature, AdultWeekend ticket price, you see quite a few reasonable correlations. fastQuads stands out, along with Runs and Snow Making\_ac. The last one is interesting. Visitors would seem to value more guaranteed snow, which would cost in terms of snow making equipment, which would drive prices and costs up. Of the new features, resort\_night\_skiing\_state\_ratio seems the most correlated with ticket price. If this is true, then perhaps seizing a greater share of night skiing capacity is positive for the price a resort can charge.

As well as Runs, total\_chairs is quite well correlated with ticket price. This is plausible; the more runs you have, the more chairs you'd need to ferry people to them! Interestingly, they may count for more than the total skiable terrain area. For sure, the total skiable terrain area is not as useful as the area with snow making. People seem to put more value in guaranteed snow cover rather than more variable terrain area.

The vertical drop seems to be a selling point that raises ticket prices as well.

We then plotted each variable against ticket price to see the relation a bit more clearly, and unhide any potential correlations we may have missed.

Now that we know how we want to think about the attributes, we start modeling. Taking out our resort, first we fill in the NA values, trying both median and mean values. First we split the data into four dataframes -- One with 70% of the prices, one with 30% of the prices, One with 70% of the other data, and one with 30% of the other data. We inserted some values for the NA values, and then scaled all of the values in the other data dataframes so that we had the ability to compare values on different units. Then we ran the regression model on the train 70% of the data to see how the variables corresponded to the price output. Using that model, we looked at the test 30% of the other data and made predictions for the prices that would come out of those values. Then we ran the error statistics on what they actually are vs what was predicted.

Once we had done that, we used the pre-built version of this from scikitlearn. Then, because we weren't sure which attributes actually helped or hurt, we tried to select different amounts of attributes, which wasn't very successful. Then we learned about folding and applied the scikitlearn model that can apply folding. We are likely to use the random forest model going forward.

Big Mountain Resort modelled price is $92.39, actual price is $81.00.

Even with the expected mean absolute error of $10.44, this suggests there is room for an increase.

If we add a run 150 feet lower and add a chair lift to support it:

This scenario increases support for ticket price by $16.28

Over the season, this could be expected to amount to $28486111

This seems to be the best scenario, and does not require any extra snow making capability in order to make it more viable.