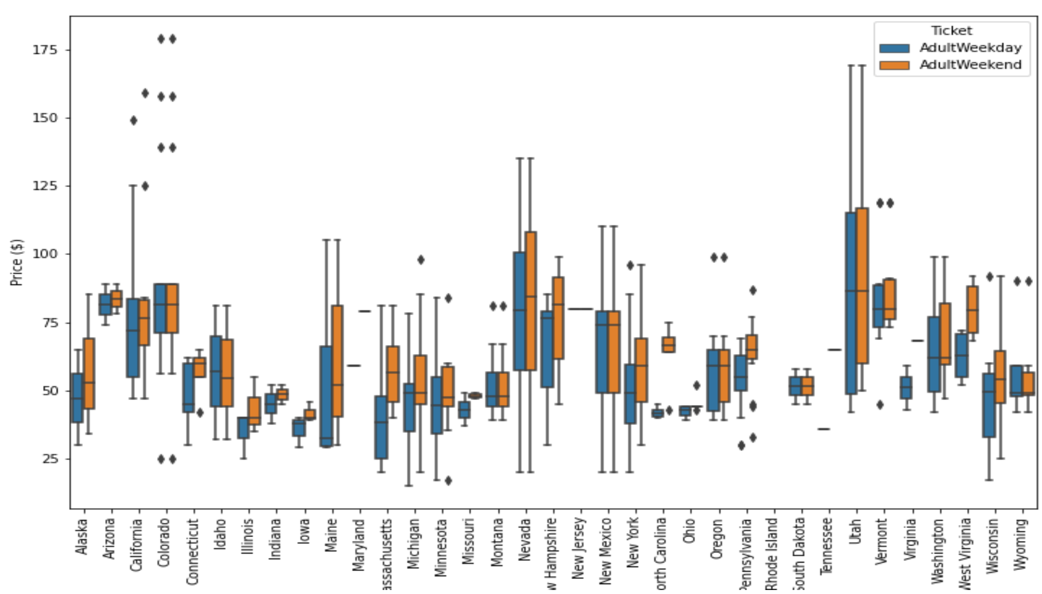
Report on recommendations for Big Mountain Resort

In this project we investigated several possible options for increasing the profit for Big Mountain Resort. More specifically, we want to understand whether the existing facilities and/or adding new facilities can support a higher ticket price.

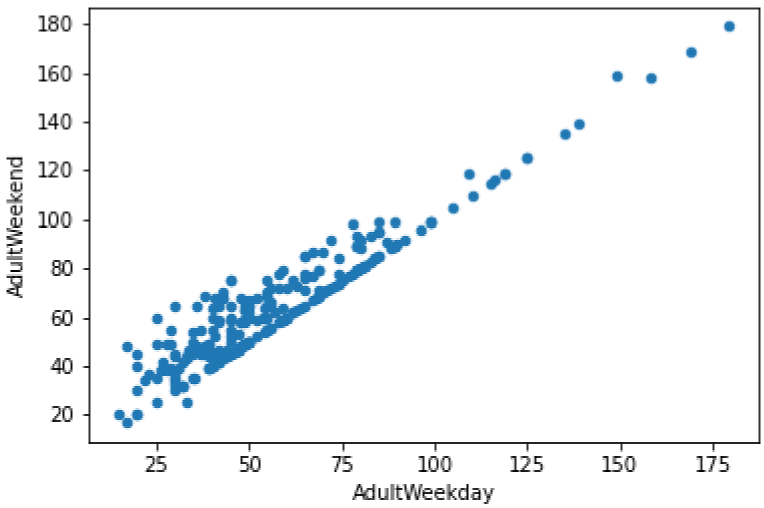
Details about our analysis:

First, we have two different kinds of ticket prices, one is weekday price and the other is weekend price. We noticed that the distributions of the two ticket prices are similar for most states. For many resorts they are actually equal. Because weekend price has the least number of missing values, we decide to model weekend price.

Distribution of weekday price and weekend price for states:

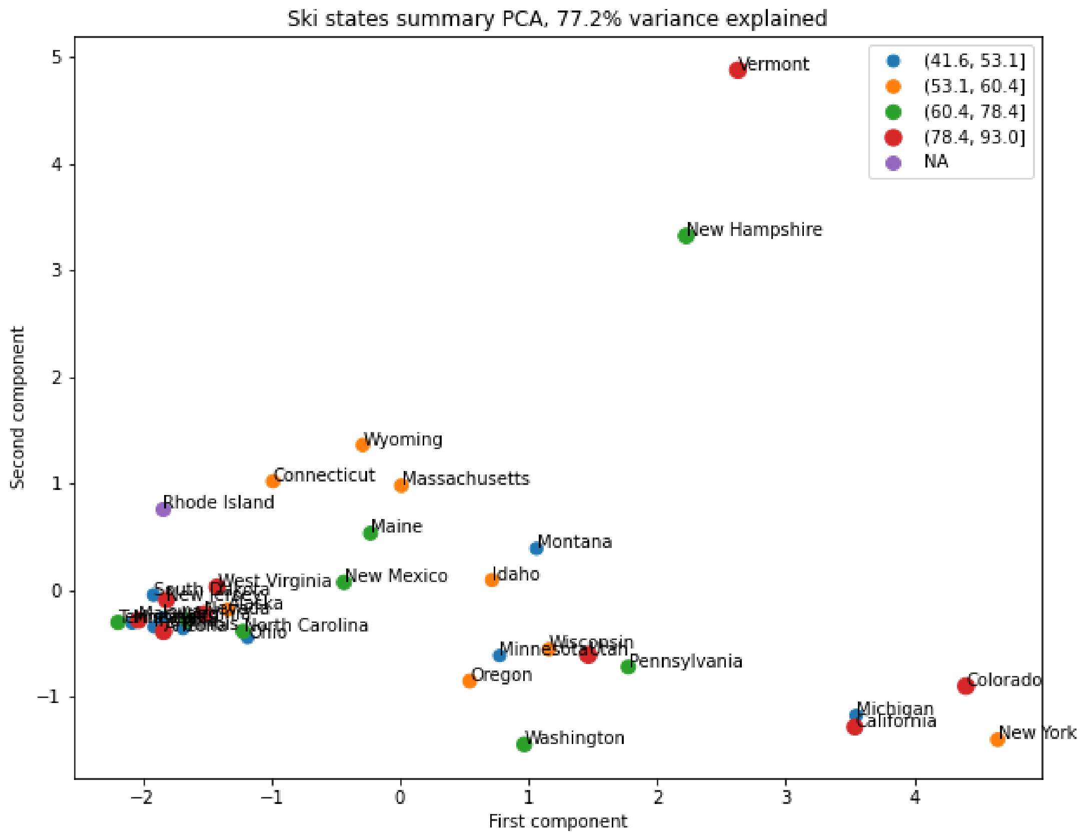


Plot comparison between weekday and weekend ticket price:

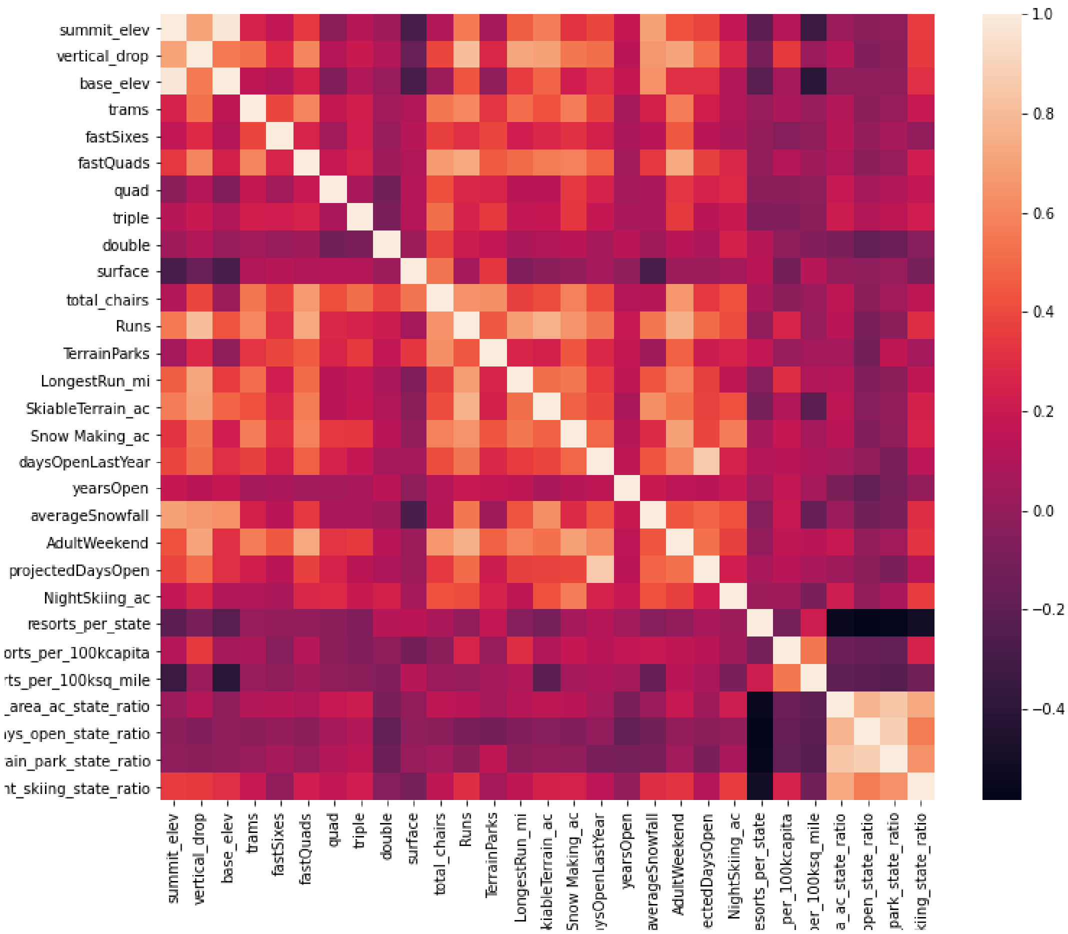


Next we investigated the relationship between state and ticket price. Because the features such as population/area/number of resorts differ significantly between states, to make comparison meaningful we obtained information for each state from Internet. From this we generated resort density features, such as number of resorts per population and number of resorts per area. Then we use PCA method to all features to project our data into 2d space, we make scatter plots using the first two components for all states, take the price level into account. We cannot see a clear pattern from this. Besides, the distribution of the mean ticket price for different states also does not reveal a clear group pattern. Based on this observation, we will treat all states equally in this project.

Plot for PCA results:



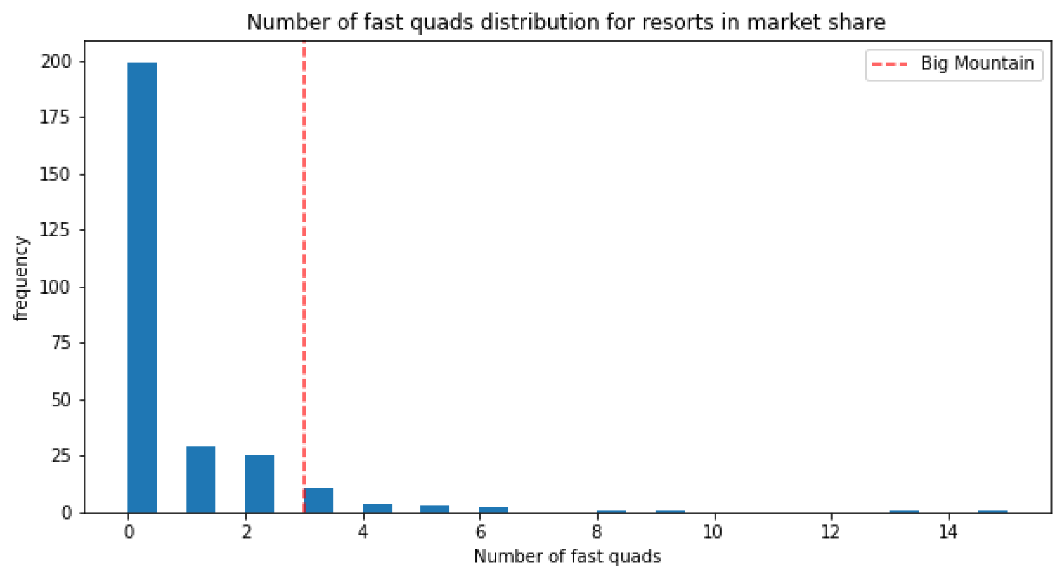
Correlation matrix:

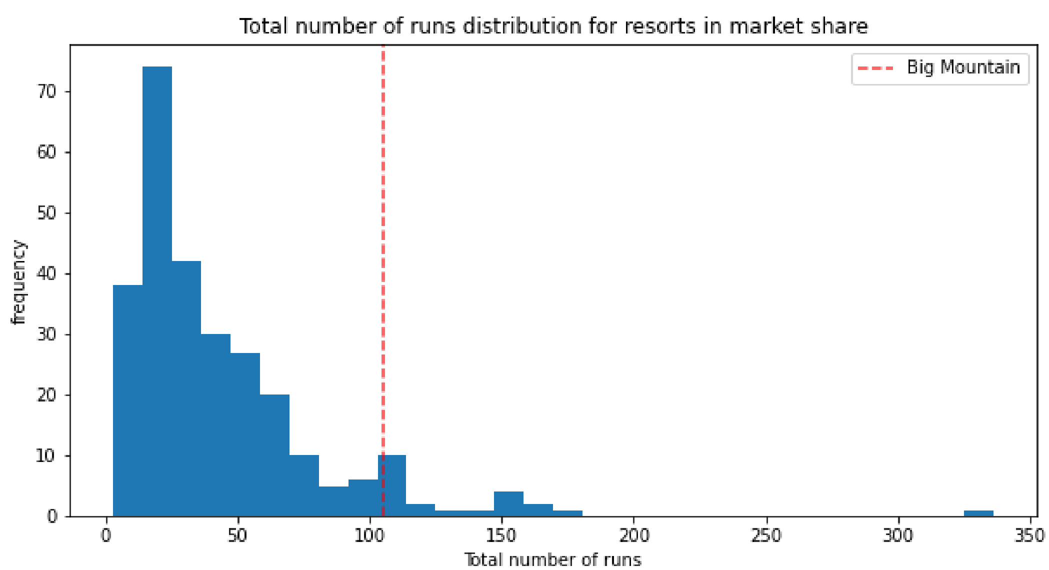


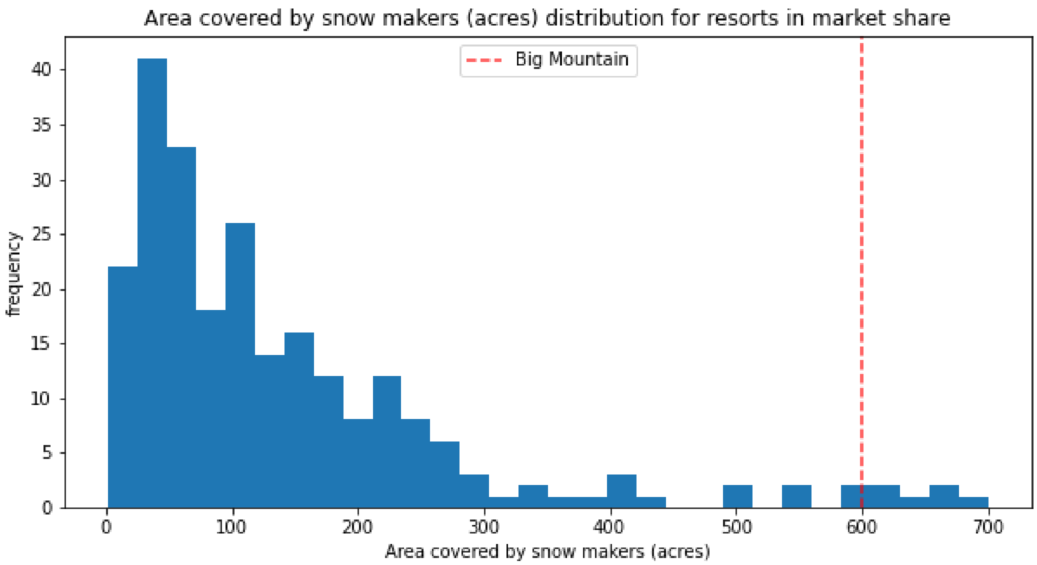
Next we combine the state data with the ski data we obtained last time. Using this we also generated some new features, such as the ratio between skiable area of a resort and the total skiable area of the state. Then we looked at the correlation matrix for all features. We find that vertical_drop, fastQuads, Runs and total_chairs show significant positive correlation with ticket price, so those features might be important in modeling the ticket price.

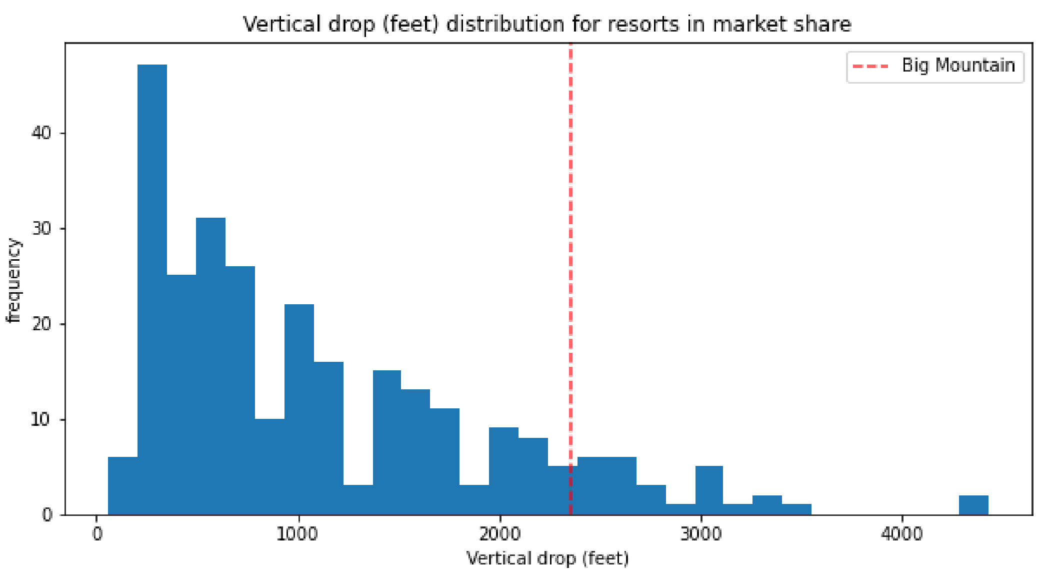
Next we tried linear regression and random forest model to predict ticket price from our features. The primary metric we use to evaluate the performance of the model is MAE. For the linear model, we find that if we feed all the features into the model, we find that the model is overfitting. To reduce overfitting,we tried feature selection, we use grid search plus cross validation for this purpose. Then we discovered that we should choose the number of features to be 8. Then we use a similar procedure as above to determine the optimal hyperparameters for the random forest model. Lastly we compare the performance of the two models and select the best one. The mean and standard deviation of MAE for the linear model in cross validation are 10.50/1.62, the MAE on the testing set is 11.79. For the random forest mode, the mean and standard deviation of MAE in cross validation are 9.77/1.38, the MAE on the testing set is 9.64. From this we conclude that the random forest model has better performance than linear model.

Next we make use of the random forest model to understand relationships between facilities and ticket price. We find that the number of fast four person chairs(fastQuads), the number of runs, total area covered by the snow making machine and the vertical change in elevation from summit to base have the most influence on the ticket price. Big Mountain is already fairly high on the league chart for these facilities, so adding new facilities is unlikely to result in a significant increase in profit.

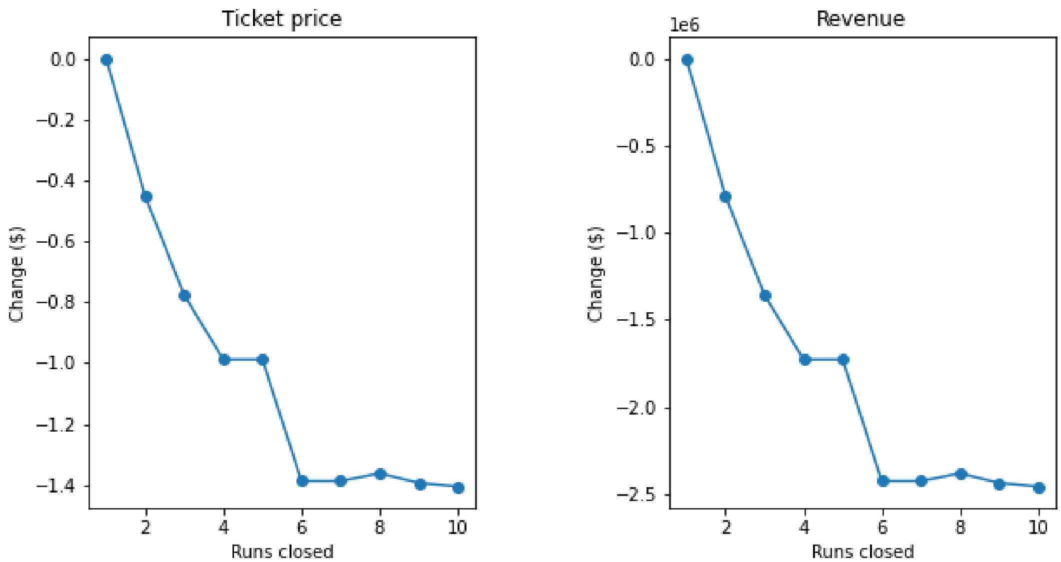








We evaluated several proposed solutions. First, we evaluated the option of adding a run, increasing the vertical drop by 150 feet, and installing an additional chair lift, our model suggests an increase by $1.08 in ticket price, which amounts to an increase of $1895250 in raw profit. Given that the recently installed chairlift already cost $1540000, this strategy is unlikely to increase the net profit. We also considered adding a run, increasing the vertical drop by 150 feet, installing an additional chair lift and adding 2 acres of snow making. The result from our model is similar to the previous one, with more operating cost, so this is not a good option. We also considered increasing the longest run by .2 miles and guaranteeing its snow coverage by adding 4 acres of snow making capability. Our model suggests this won't increase the ticket price either. Our last option is to close up to 10 of the least used runs. Our model suggests that doing this will reduce the ticket price, but if we take the reduce in operating cost into account, then this can result in an increase in net profit, especially if we do not close more than 5 runs.



Our recommendation:

Based on our analysis above, our suggestion is to gradually close up to five the least used runs while fixing the ticket price. Each time we close a run, we can wait for sometime to evaluate people's response. If we see an increase in net profit, then we can consider closing more runs.

Conclusion and future direction:

Our model indicates that the Big Mountain Resort is unlikely to benefit from adding new facilities. But there are also limitations in our dataset. First, since the net profit depends on both ticket price and the number of tourists, our analysis will be more accurate if we can get the number of tourists per year for each resort. Next, it is also better to know the operating cost for each resort, this will help to better analyze the net profit. Lastly, the ticket price may also depend on the economical situation for states, so we might also need to take the average annual income for each state into account, for instance. With these additional information, we can perform a more comprehensive analysis.