**Guided Capstone Project Summary – Big Mountain Resort**

Big Mountain Resort (BMR) is a ski resort located in Montana, with views of Glacier National Park and Flathead National Forest and access to 105 trails. Every year about 350,000 people ski or snowboard at BMR.

BMR’s past approach to ticket pricing has been to charge a premium above the average price of resorts in the same market. However, this approach has limitations and there may be room for optimization, particularly as far as capitalizing on its facilities. BMR’s leadership would also like know which facilities should be developed for maximal return.

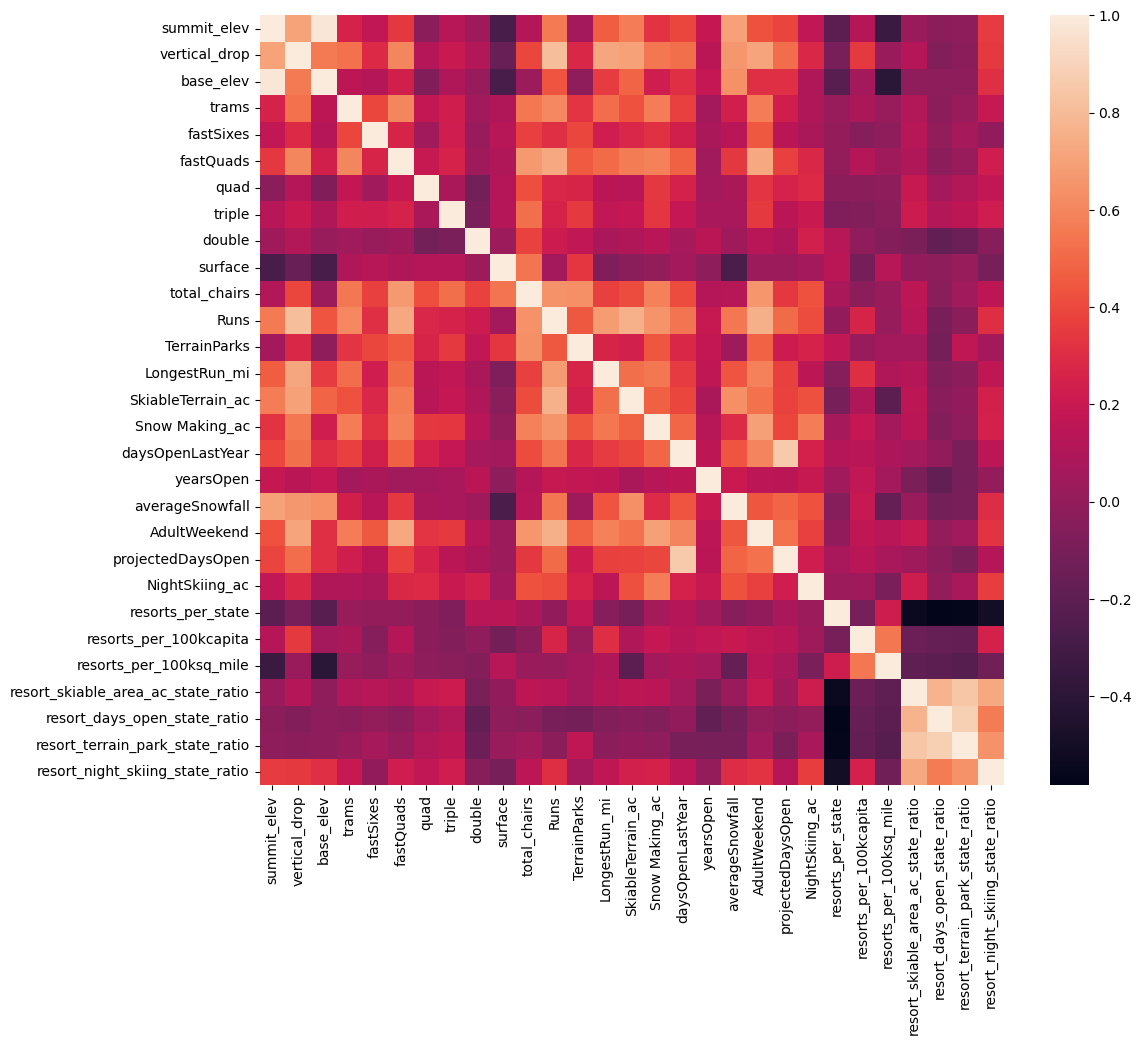
Therefore, the goal of the goal of the project was to identify which facility features have the largest effect on ticket prices based on the market share data provided. To do this we built a predictive model for ticket pricing using these features. We were provided with a CSV file containing market share data from across the country.

**Data Wrangling:**

The first step in this project was Data Wrangling, in which we organized, cleaned and made sure our data was well defined. Here, we quantified and imputed missing values, looked at feature distributions for outliers, and dropped irrelevant data. We then imported some external population and area data for each state. After examining some options for our Target Feature, we settled on the AdultWeekday ticket price as our target variable as it had fewer missing values than the other ticket price option.

**Exploratory Data Analysis (EDA), Model Preprocessing & Feature Engineering:**

Next, we performed some EDA to get a sense of some of the basic characteristics of the data. First, we looked at whether there was a relationship between state and average ticket price, and found that there wasn’t one, and therefore decided not to use state as a relevant factor for determining ticket price. However, we did find some patterns from the geographic analysis suggesting that resort density (per capita and per square mile) may be relevant factors for modeling; so we calculated those ratios and added them to the dataframe. We then did PCA analysis to visualize how linear combinations of the original features contribute to the variance, and whether there was any pattern related to ticket price. We determined that there wasn’t and moved on to looking at other factors.



Using the heatmap above, we plotted the feature correlations to examine relationships between them. This was useful in identifying some key features on which to focus our model, for example, fastQuads, runs, total\_chairs, vertical drop, Snow\_Making\_ac, and resort\_night\_skiing\_state\_ratio. Finally, we learned that some features may be related to each other, for example the resort terrain and the resort skiable area are likely to be somewhat dependent.

**Algorithm Exploration & Evaluation Metrics**

In this step, we split the data into training and test sets, made an initial guess of ticket price using the mean of the Adult Weekend ticket price, and then measured the mean absolute error and mean squared error. We found that ticket prices would be off by about $19 or $24 on average, using this very basic algorithm.

Next we scaled the data, built a linear regression model, and then assessed the model’s performance. We used cross validation to ensure the model could generalize to new data and found the optimal number of features to be 8, and the feature with the largest positive effect on ticket price to be vertical drop. We also found that the simple linear regression model could account for 81.8% of the variance on the training set and over 72.1% on the test set. We could expect to get a ticket price within about $9 of the actual price using linear regression.

Next we tried the Random Forest Model, and explored varying some parameters of the model, such as number of trees, scaling, and imputing methods. We found that imputing using the median improved performance, but scaling did not. Estimated performance via cross validation was about 71%. Importantly, we found the top four features were fastQuads, Runs, Snow Making acreage and vertical drop in determining ticket price. (See figure below.)A graph with blue and white text

AI-generated content may be incorrect.

**Winning Model and Modeling Scenarios**

Once we had two adequately performing models, we had to choose between them. We noted that the Random Forest (RF) model had a lower cross-validation mean absolute error by almost $1 and also exhibited less variability. Therefore, we chose to proceed with the RF model. We also evaluated whether we had a sufficient amount of data in order to achieve our goal, and determined that the CV score plateaued after about 50 data points (see figure below). Therefore no further data collection was necessary.

A graph with a line

AI-generated content may be incorrect.

Big Mountain Resort currently charges $81.00 for an Adult Weekend ticket. However, based on the resort’s current facilities, our model suggests that ticket price could be increased to $95.87, with a mean absolute error of $10.39. Therefore, there is room for an increase.

The increase in ticket price can be justified by the fact that Big Mountain Resort falls on the high end of almost all of feature distributions that were identified as being important in setting ticket prices, including vertical drop, snow making acreage, chairs, fast quads, runs, and skiable terrain.

For this project, BMR leadership was considering four possible scenarios for increasing revenue. We explored these scenarios in terms of their effect on ticket prices and revenue, as follows:

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

**Result:** No effect on either ticket prices or revenue by closing 1 run. A decrease in ticket price and revenue by closing 2 runs, and then another drop by closing 3 runs. There is no difference in closing 3, 4 or 5 runs. Significant drops if closing 6 or more runs. (See figures below.)

A graph of a price

AI-generated content may be incorrect.

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

**Result**: This scenario supports increasing ticket prices by $1.99 and revenues of $3.47million.

1. Same as number 2, but add 2 acres of snow making coverage.

**Result**: Adding the snow making coverage has no effect on ticket prices.

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

**Result**: Increasing the longest run has no effect on ticket prices.

Therefore, it would appear that Scenario #2 would be the best option for increasing ticket prices, provided that the chair lift has already been installed and the cost of installation does not need to be accounted for. However, the cost of operating and maintaining the additional chair lift does need to be considered. In order for this scenario to be justified, the annual operating/maintenance costs of the chair lift must be less than $3.47 million. Or, we could add that cost to the ticket price, as follows: $OpCost/(5\*350,000).

**Recommendations & Conclusion**

In conclusion, there are indeed opportunities for increasing revenue at Big Mountain Resort. We recommend the following, in order of importance/urgency:

1. Ticket prices can immediately be increased to a maximum of $95.87, based on existing facilities and comparison to the rest of the market.
2. Close one run in order to cut costs, without any reduction in ticket prices or revenue.
3. Increase the overall vertical drop by installing one chair lift, but only if the annual operating and maintenance cost for that chair lift will be less than $OpCost/$1.75mil. Ticket prices can then be raised an additional $1.99.
4. In the longer term, BMR should look for additional opportunities to increase the vertical drop, since that factor seems to have the greatest effect on ticket pricing.

**Future Scope of Work**

This model has the potential to be improved as more data becomes available, and will likely be useful in the future for exploring other scenarios. Therefore, I would rewrite it to make sure its generalized (eg. Important values such as visitors per year are parameterized), and then turn it into a standalone application with a GUI that the general staff/analysts can use.