# Problem Statement

The purpose of this analysis is to develop a machine learning prediction model to estimate ticket pricing for Big Mountain Ski Resort (Big Mountain) based on the pricing and amenities of the resort's competitors.

The analysis is subject to the following assumptions:

1. Competitor resort pricing is based on the free market value
2. Ticket pricing is dependent on included resort facilities.

# Data Wrangling

The dataframe, provided by the Client, consisted of categorical data for 330 resorts. The data was examined for missing values and outliers which may impact or skew the results of the model. Therefore, several items were either updated or removed from the dataset:

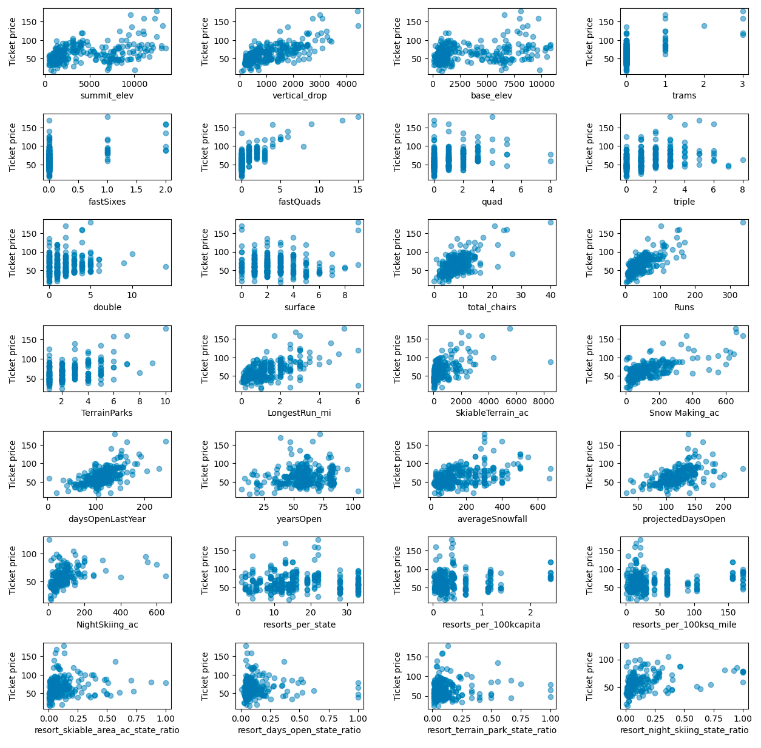
1. Approximately 16% of resorts were missing ticket price data, with approximately 14% of the resorts missing both weekend and weekday pricing. The resorts missing both weekend and weekday pricing were removed from the analysis. The missing ticket price values within the remaining resorts were estimated using the median value.
2. Outlier values were identified in two categories: skiable terrain (Silverton Mountain) and years of operation (Howelsen Hill). Silverton Mountain in Colorado was reported to have a skiable area of approximately 26,816 acres, which is significantly larger than the competitor resorts. Upon further investigation, the skiable area was corrected to 1819 acres, per the Silverton Mountain website. Additionally, Howelsen Hill in Colorado was reported to have been open for 2019 years. Due to the publication of the data (2019) this was likely the year the resort opened. However, it may also mean the resort was not open at the time the data was collected. The resort was removed from the analysis due to the lack of information regarding the timeline,

The number of resorts remaining in the dataframe is 277.

# Exploratory Analysis

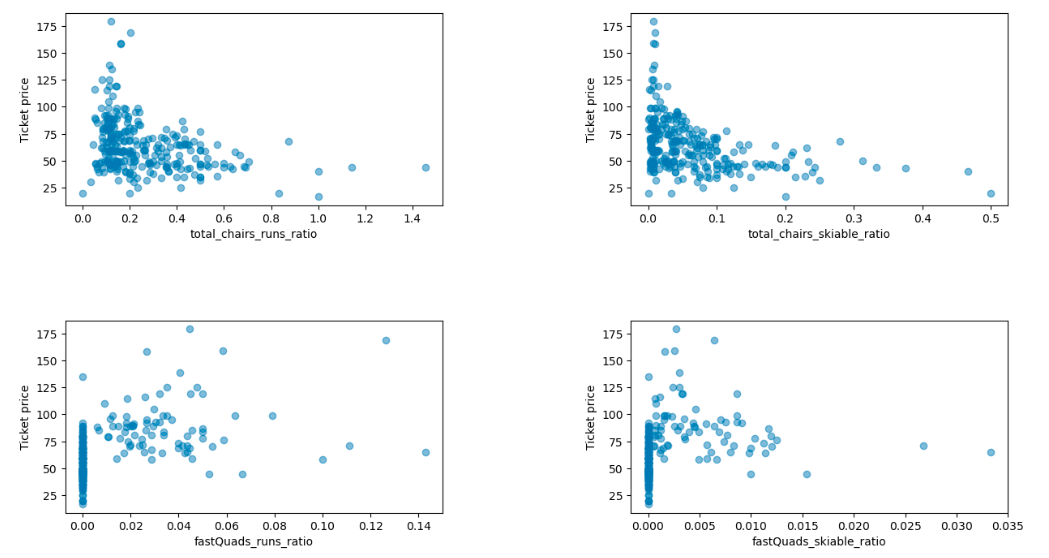
A principle components analysis (PCA) was completed to identify linear combinations of uncorrelated features and variance. The results were used to develop a series of scatterplots to display the relationships between ticket price and remaining variables. Refer to **Figure 1**.

The results show a strong positive relationship between the ticket price and vertical drop along with a positive relationship between ticket price and fastQuads, runs, and total chairs.Additionally, the resorts per 100k capita plot shows when the number of resorts per capita is low, the ticket price is variable and ticket prices tend to increase as the number of resorts per capita increase.



**Figure 1: Numerical Features Vs. Ticket Price**

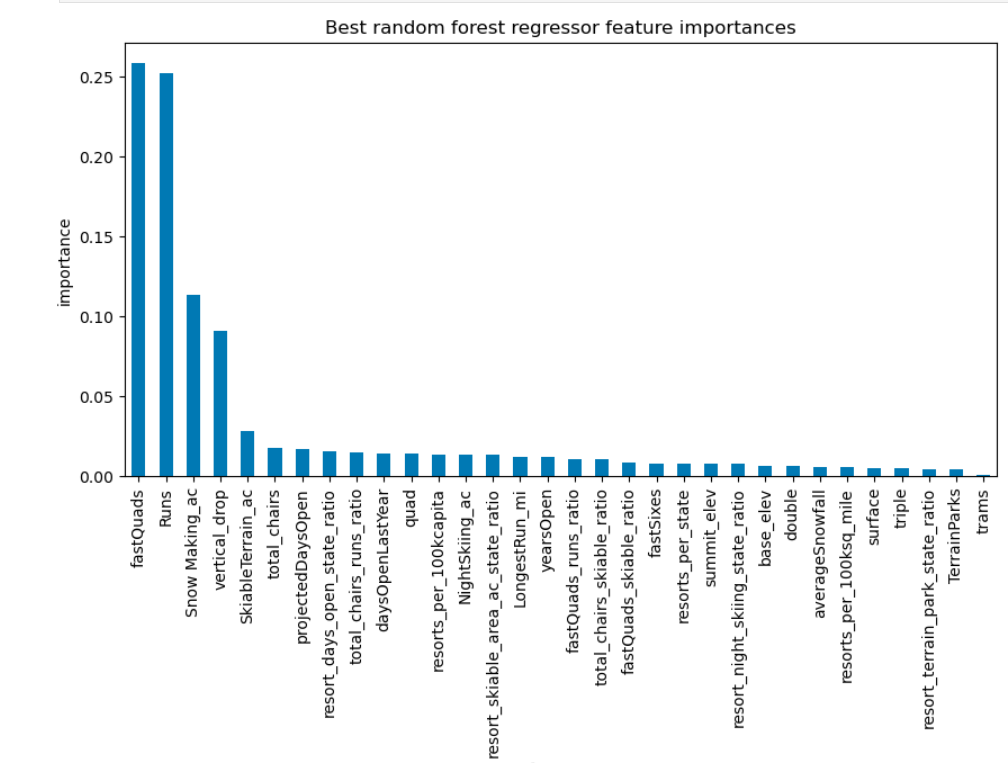
Additional plots were developed to determine how transportation may impact the ticket prices at resorts. This was accomplished by comparing ticket prices to the ratio of chairs:runs and chairs:skiable terrain (**Figure 2**). The results show a negative correlation, potentially indicating an increase in skiers on the mountain may decrease the prices.

****

**Figure 2: Chairs Vs. Ticket Price**

# Preprocessing and Training

A Random Forest Regression Model was employed to estimate the ticket price for Big Mountain. A Random Forest Regression Model is a machine learning predictive model which combines independent predictions through a weighted average to reduce overfitting and improve accuracy. Therefore, the model was trained by partitioning the original data into numerous combinations of the dataset, consisting of a training subset (i.e. 70% of the database) and validation subset (i.e. 30% of the data) to determine the best fit of weighted data. The optimal combination of hyperparameters (data subsets), presented in **Figure 3**, was selected to develop the best fit model.



**Figure 3: Random Forest Regression Model Features of Importance**

The cross-validation score, or performance ranking of the model, was determined and is presented in **Figure 4**. The results show the model rapidly improves as more data points are added, then levels off at a sample size of approximately 50. Therefore, additional data collection is not required to improve the model results.



**Figure 4: Cross Validation Score**

# Conclusions

The predictive model determined the fair market price of an adult weekend ticket to Big Mountain falls within the range of $85.48 and $106.26. Thereby, suggesting a price increase between $4.48 and $25.26.

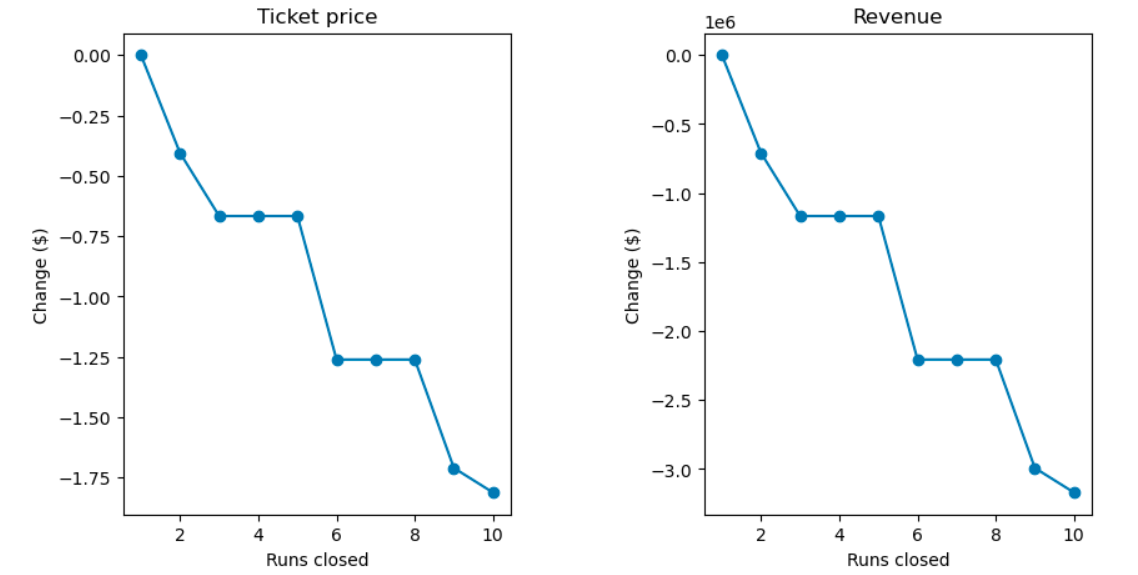
Recall the results are impacted by the assumptions and limitations of the predictive model. For example, the model assumes the competitor resorts are charging a fair market price dependent on the facilities examined within the dataframe. However, it is reasonable to suggest the competitor resorts may be under- or overcharging, and/or additional factors may impact the fair market value of the ticket. Therefore, additional market research is recommended to lower the uncertainty of the model and better understand the ticket pricing model employed by competitor resorts.

## Alternatives Analysis

Per the request of the client, we examined four cost saving alternative scenarios, listed below:

1. Closing up to 10 of the least used runs;
2. Increasing the vertical drop and installing an additional chair lift;
3. Increasing the vertical drop, constructing an additional chair lift, and adding an additional 2 acres of snow making cover; and
4. Increasing the longest run by 0.2 mile to boast 3.5 miles length.

The financial impact of Alternative 1, closing up to 10 of the least used runs, is presented in **Figure 5**. Note, closing one run will not impact pricing. Closing 2 runs will reduce seasonal revenue by approximately 700,000 dollars. Closing 3 to 5 runs will reduce the seasonal revenue by approximately $2,300,000. And, closing additional runs leads to a large drop in ticket prices and revenue.



**Figure 5: Ticket Price and Revenue Vs. Run Closures**

Alternative 2, increasing the vertical drop and construction of an additional chair lift will increase seasonal revenue by approximately $3,500,000. However, there will be no additional revenue with the addition of 2 acres of snow making cover (Alternative 3). Lastly, increasing the longest run by 0.2 miles to boast 3.5 miles length (Alternative 4) will not impact revenue. However, this is likely due to the limitations of the model, as the run length feature has a low impact within the model design.

# Recommendations

Based on the results of the predictive model, there are several options to increase revenue and/or lower operational costs at the resort. Increasing ticket prices to better match market value will increase profitability. Additionally, Alternatives 1, 2, and 4 are potential solutions to increasing revenue. However further analysis may be required.

Closing runs will reduce seasonal revenue, and lower operation costs. Increasing the vertical drop and adding an additional chair lift will increase revenue, and require construction costs and increase operation costs. Additionally, extending the longest run may impact revenue. However, the model will need to be adjusted to reflect the actual influence of each facility on ticket pricing.

We recommend collecting the following data and updating the predictive model, accordingly. This will decrease the error of the results, and likely lower the range of the suggested ticket costs. Additionally, this information will be used to develop a hierarchy, or weight, of each facility on ticket pricing, and support a cost/benefits analysis which may be used to determine facility construction or closures.

1. Additional amenities and its cost/usage/profitability (ex. lodging, parking, equipment rentals, dinning, classes, ski route difficulty level/variability, etc.)
2. Number of visitors of the resort and given facility
3. Accessibility (ex. parking, transportation, distance to population/travel centers, etc.)
4. Operational costs for resort’s facilities