# Guided Capstone Project Report

## Phong Pham

**Introduction**

Big Mountain Resort has recently installed an additional chairlift to help increase the distribution of visitors across the mountain and this increases their operating costs by $1,540,000 this season. The resort wants some guidance on how to select a better value for their ticket price. They are also considering changes that they hope will either cut costs without undermining the ticket price or will support an even higher ticket price. The resort management wants to find data-driven solutions to make up for the new expenses and keep the business profitable, based on data from other U.S. ski resort in the same market segment. Analysis on other resorts’ trails, amenities and operations can help Big Mountain come up with a better price prediction model than using the average price of all other resorts.

**Dataset**

The data contains details about 330 U.S. ski resorts, such as the number of chairlifts, trails, vertical drop, and daily operations. The number of fast eight chairlifts was missing from more than half of the resorts, and all but one values were zero, so we dropped the whole column as it did not convey essential information. We also decided to drop the resorts that have no price data for neither weekday nor weekend, and use weekend ticket prices to predict Big Mountain new price as they were missing fewer values when compared to weekday prices. This left us with data of 277 ski resorts.

**Preprocessing**

We start by excluding Big Mountain record in the dataset, then split the data into a training set and a testing set with a 70/30 ratio. If the average ticket price was used to predict the new ticket price for Big Mountain Resort, the value would be off by around $19.

I created a linear model by imputing missing values using either the median or mean values of all fields, scaling the data to put all the data on a consistent scale, selecting the k best features and training the model. With this model, the price prediction error was estimated to be within around $10 of the real price. Vertical drop was found to be the best predictor of the ticket price, followed by the area covered by snow making equipment. On the other end, the number of trams and the area of skiable terrain is negatively associated with ticket price.

A random forest regressor was also used, and we found that the number of fast quads and runs had the highest positive effects on ticket price. These were followed by vertical drop and area covered by snow making equipment, which also were the best predictors using linear regression model.

Chart

Description automatically generated

The mean absolute error of the random forest regressor cross-validation was smaller by almost $1 than the linear regression result, and there was also less variability in the forest regressor model. Therefore, I decided to use the forest regressor model to predict the price for Big Mountain.

**Modelling results**

After using forest regressor model, the ticket price was calculated to be $95.37, compared to the current price of $81.

Big Mountain is comparatively high up the league table among the resorts for current weekend ticket price.

Chart, histogram

Description automatically generated

Big Mountain is also ranked highly in the number of fast quads, runs, the area covered by snow making equipment, and vertical drop. The resort has shortlisted the following 4 options to either cut costs or increase the support of ticket prices:

* Close down up to 10 of the least used runs: This would result in a drop in the support of ticket prices and revenue if 2 or more runs were closed.

Chart, line chart

Description automatically generated

* Increase the vertical drop by adding a run to a point 150 feet lower down and install an additional chair lift: This scenario increases support for ticket price by $1.99. Over the season, this could be expected to amount to $3,474,638
* Same as the above scenario but adding 2 acres of snow making cover: This would make no difference in the ticket price whatsoever.
* Increase the longest run by 0.2 mile to boast 3.5 miles length, requiring an additional snow making coverage of 4 acres: This would also make no difference in ticket prices.

Therefore, If the resort increases the vertical drop by 150 feet and installs an additional chair lift, the support for ticket price would be increased by $1.99, which would amount to a $3,474,638 revenue increase over the season, on the basis of each visitor buying 5 day tickets on average. This would help the resort cover the additional operating cost of the new chair lift, which amounts to $1,540,000, and make an additional profit of $1,934,638. Therefore, I would recommend the business to implement scenario 2. If Big Mountain wants to test the scenario with run closures, I would suggest that they close no more than 5 runs to able to cover the new chair lift operating cost, as the decrease in revenues when closing 6-10 runs is projected to be greater than $1,934,638. It means that they would not have enough revenue to cover the new lift operation expenses if more than 6 runs are closed.

**Future**

The model was limited by not having any data on other operation costs. It would be better to have data on the expenses to increase vertical drop, snowmaking area, add a run, or even close down any runs. The modeled price was much higher than its current price possibly because Big Mountain ticket price is the highest in the state of Montana, and the competition within the state might have made the management adjust the price lower. If the executives find the model successful, they can implement any combination of the 4 scenarios based on the modeling results. The analysis might be made available to business analysts via a website where they can explore by manually selecting any combination of parameters to see the projected results.