*Janis Navarro*

*Springboard Data Science Career Track | Guided Capstone Project Report*

*Big mountain resort*

What if any opportunities exist to efficiently develop and implement a new ticket pricing strategy to maximize capitalization and offset operational costs?

**Introduction:**

Big Mountain Resort is a ski resort located in Montana. The resort offers spectacular views of Glacier National Park and Flathead National Forest, with access to 105 trails. They service 11 lifts, 2 T-bars, and 1 magic carpet for novice skiers. The base elevation is 4,464 feet and the summit is 6,817 feet with the longest run at 3.3 miles in length.

Big Mountain Resort recently installed a new chair lift to increase the distribution of its visitors across the mountain. The new chair lift increased operating costs by $1,540,000 this season. The increase in operating cost has led the business to rethink their ticket pricing strategy of charging at premium above the average price of resorts in its market segment.

**Business Problem:**

Big Mountain Resort wants to come up with a pricing model for ski resort tickets in our market segment. They want to maximize its returns relative to its position in the market and to have a strong sense of what facilities matter most to visitors and which are the most likely they are willing to pay more for. The new pricing strategy will take into consideration data gathered from ski resorts nationwide.

How can we create a pricing model that can provide guidance for Big Mountain’s pricing and future facility investment plans?

**Data Wrangling:**

Audit of our data displayed that the dataset includes several important values such as weekday and weekend ticket price, number of lift chairs, total vertical drop, total number of runs for each resort. The data was first reviewed for ticket prices. We ran into that the dataset includes two sets of different ticket price values for Weekday labeled AdultWeekday and Weekend labeled AdultWeekend. We reviewed AdultWeekday vs AdultWeekend prices for:

What relationship is there between these weekday and weekend prices? Most States had the same price for Weekday and Weekend as visualized in this chart:

Chart, scatter chart

Description automatically generated

Weekend prices had the least missing values of the two, so we removed the AdultWeekday prices. In addition to AdultWeekday, the fastEight column was also removed because half the values were missing and all, but the others are the value of zero. There were also approximately 14% of the data had both prices missing and were dropped. An outlier value was imputed with a correct value. Additional rows that had missing values were removed as well. We also obtained state population and size data. After this left us with 277 rows of the original 330 and 25 data columns of the original 27 columns and we identified we will predict the adult weekend ticket price.

**Exploratory Data Analysis:**

To find relationships, trends, patters, we explored the data. The first pattern explored was the relationship between the total number of resorts by population vs the total number of resorts by area. This data did not yield usable results for us, but it helped lead the exploratory process. We then explored the relationship between states. There wasn’t any pattern, so I decided to treat all states equally. The next relationship through our EDA was the relationship between the components like vertical drop, years open, or skiable areas versus the price in each state. This was completed through a Principal Cumulative Analysis (PCA). The PCA showed in the first two components, there is a spread of states across the first component. The first two components account for over 75% of the variance, and the first four for over 95%. I was then able to explore the resort-level data in more detail. To gain a high-level view of relationships amongst the features a heatmap was created to visualize it as follows:

Chart

Description automatically generated

Turning attention to our target of AdultWeekend ticket price, we see fastQuads, Runs, Snow Making\_ac and resort\_night\_skiing\_state\_ratio are the most correlated with ticket price. Total\_chairs and the vertical drop seems to be quite correlated with ticket prices as well. We will use these features to build a model for ticket price.

**Pre-Processing and Training Data:**

Taking what I learned from the EDA process, we began by seeing how good the average price is for pricing. We first tested the mean, which did not yield good results giving us a value at $63.81. We also measured the Mean Absolute Error as it is arguably the most intuitive of all the metrics. This told us that on average, we might expect to be off by around $19 if we guessed ticket price based on an average of known values. This was much too large. So we moved on and performed a linear regression using the median between results. The Mean Absolute Error then measured on average you’d expect to estimate a ticket price within $9 or so of the real price. I then moved to create a data pipeline in order to efficiently produce comparable results. These results suggested that vertical drop is Big Mountain’s biggest positive feature. Which is comparable to the results from our EDA.

The next regressor was a Random Forest Model that identified imputing the median value helps with the Mean Absolute Error of our four components. I also included the vertical drop to the Random Forest Model. The results were astonishing as the model has a lower cross-validation mean absolute error by almost $1. And it exhibited less variability. Here is a chart of the best random forest regressor with the highest importance features in order:

Chart, histogram

Description automatically generated

**Modeling:**

Having chosen the top components from the Best Random Forest Regressor Feature Importances mentioned in the Pre-Processing and Training Data, we also used features that came up as important in the modeling combined to include:

* Vertical Drop
* Snow Making Area
* Total Number of Chairs
* Fast Quads
* Total number of Runs
* Longest Run in Miles
* Trams
* Skiable Terrain Area in Acres

To review the ticket price, we needed to see where Big Mountain Ski Resort ranked within all the above-mentioned features. Big Mountain Resort is represented by the dashed red line as follows:

**Vertical Drop**

Chart, histogram

Description automatically generated

**Snow Making Area**

Chart, histogram

Description automatically generated

**Total Number of Chairs**

Chart, histogram

Description automatically generated

**Fast Quads**

**Graphical user interface

Description automatically generated with low confidence**

**Total Number of Runs**

**Chart, histogram

Description automatically generated**

**Longest Run in Miles**

**Chart, histogram

Description automatically generated**

**Trams**

**Chart

Description automatically generated**

**Skiable Terrain Area in Acres**

**Chart, histogram

Description automatically generated**

All these present that Big Mountain Ski Resort ranks well above average or high in each category. Apart from Trams, which Big Mountain Ski Resort matched the majority at zero number of trams.

From these we can gather that Big Mountain Ski Resort operates in a market where people pay more for certain facilities, and we are in great place to capitalize on that with our ticket price.

Our expected Big Mountain Ticket Price from the model is $95.87, which is way above the current price of $81.00. Even with the expected mean absolute error of $10.39, the model suggests there is room to increase our ticket price.

**Conclusion:**

Big Mountain Ski Resort has all the top features and facilities that people are willing to pay more for. According to our data analysis, Big Mountain Ski Resort ranks above average or high on seven of the eight most important features that impact ticket pricing. The model supports an increase of at least $10.39.

The model also provided additional cost savings recommendations based on the 350,000 visitors per day. it suggests that Big Mountain Ski Resort can operate while still not having all the runs open at once as even with the 350,000 visitors per day, the runs are not being fully utilized to its full capacity. It specifically suggests that Big Mountain Ski Resort can keep 3 to 5 runs closed without an impactful drop in ticket revenue.

Chart, line chart

Description automatically generated

In summary, Big Mountain Ski Resort has the facilities and features that most people seek when choosing a ski resort. The data and model provided us a good improvement on current ticket prices based on the currently available data. In the future, I suggest as more information is made available for other resort amenities that people are willing to pay more on ticket price, that Big Mountain Ski Resort reevaluates the data and continues to seek opportunities for revenue growth and possible reduction on operating costs in the years to come.