Guided Capstone Project Report – Step 6

Overview  
Big Mountain ski resort located in Montana currently offers spectacular views of Glacier National Park and Flathead National Forest where every year approximately 350,000 visitors can experience this thrill. The team at Big Mountain Ski is always dedicated and focused on providing an exceptional experience for skiers and riders of all skill levels, and recently Big Mountain resort improved customer experience by installing an additional chair lift that will help increase the distribution of visitors across the mountain.

Problem Statement  
With this enhancement there are costs associated which leads us to revisit our pricing strategy. Our focus has been to determine what kind of pricing strategy Big Mountain Resort should consider offsetting the $1,540,000 annual operating expense increase, and if there is a chance where we can reduce other operating expenses to maintain and or grow our cashflow. While that is the primary issue we were trying to resolve, there is a suspicion that Big Mountain is not capitalizing on its facilities as much as it could as there is a lack of understanding to what facilities are the most important to visitors, this lack of understanding could potentially jeopardize Big Mountain’s Investment Strategy in the long run.   
  
Our Data Science team has been working diligently to provide Big Mountain Resort with a more data-driven business strategy and share our processes, winning model and price recommendations.

Data  
The Dataset we used in our analysis was provided with the support of our Database Manager, Alesha Eisen. We also used Population and area data for the US states obtained from Wikipedia.   
\**US Population and area source: 'https://simple.wikipedia.org/w/index.php?title=List\_of\_U.S.\_states&oldid=7168473'*  
  
Before we jump into the analysis stage, we needed to ensure the data was clean and relevant to our problem we are looking to resolve. The initial review of the dataset consisted of 330 rows and during the cleaning stage, we had dropped the fasteight row due to majority of the values missing, along with Adult weekend ticket prices as there were significantly higher number of missing values when compared to weekend prices, this wouldn’t have much of an impact in our analysis as all Montana AdultWeekend and AdultWeekday prices were identical. *Please refer to Fig1 and Fig 2 below.*A graph of different colored columns

Description automatically generated with medium confidence*Fig 1 shows the distribution for weekday and weekend prices in Montana are equal.*

A diagram of a person with blue dots

Description automatically generated with medium confidence

*Fig 2 – Shows there is a clear line where weekend and weekday prices are equal. Weekend prices being slightly higher than weekday prices which seem restricted to sub $100 resorts.*  
  
  
  
  
  
There were a few corrections that had to be made however it was simply user input issues, nothing that hinted towards a data migration or export error. There were additional rows dropped that did not contain any pricing data, once we had completed our cleaning stage the dataset consisted of 277 rows.

Exploratory Data Analysis

Once our data was clean it’s time for Exploratory Data Analysis, here we can focus on the pricing strategy and predict the adult weekend ticket price for ski resorts.   
  
Using the US Population and area data, we’ve explored total state area, population, resorts per state, and focused on features that would impact user’s pricing behaviour (total skiable area, total nigh skiing area, and total days open). With this we were able to scale and measure the resort density (Fig 3 and 4). We’ve also calculated the average ticket price by state (Fig 5)   
  
A graph with blue bars

Description automatically generatedA graph with blue bars

Description automatically generated  
 *< Fig 3 – shows number of resorts per 100k population*

*Fig 4 – Shows number of resorts per 100k sq mile >*

A graph of a number of blue squares

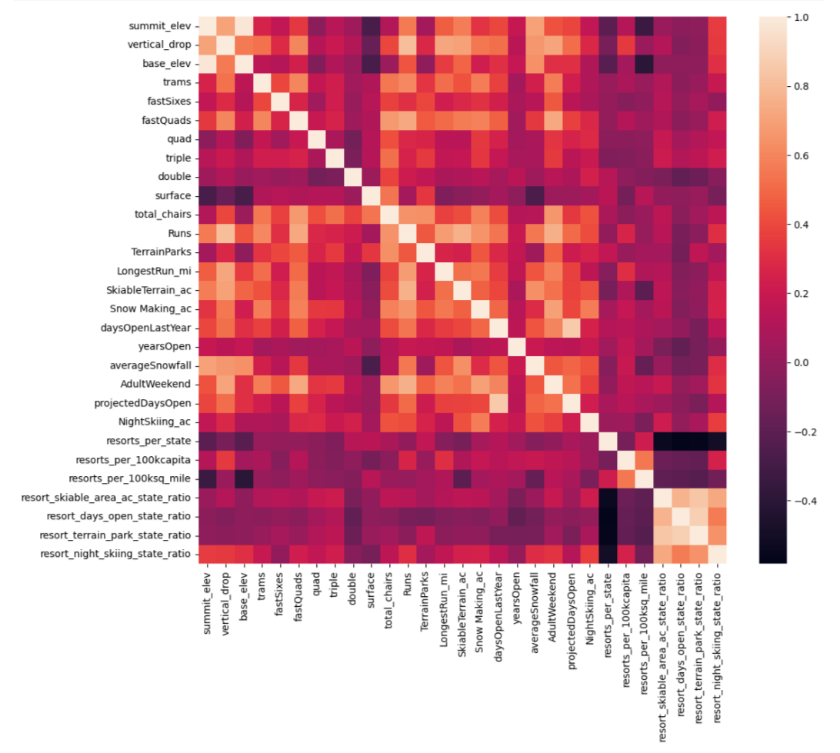
Description automatically generated

We’ve constructed some potential useful and business relevant features derived from summary statistics for each states we are concerned with. Using the resort data set we’ve explored many of these features in turn and found various trends and determined the average adult weekend price for each state (*Fig 5 on the right).*

After we merged the state summary features into the ski resort data and explored the following:

* Ratio of resort skiable area to the total state skiable area
* Ratio of resort days open to total state days open
* Ratio of resort terrain park count to total state terrain park count
* Ratio of resort night skiing rea to total state night skiing area

Once we were able to understand what share of the states’ skiing “assets” is accounted by each resort, we analyzed the correlation between the features and ticket price. (*Refer to Fig 6 below)*

We’ve found that fastQuads, Runs and Snow Making\_ac have a significant effect on ticket price.   
We’ve also found that visitors seem to value more guaranteed snow. Another feature that can be explored is resort\_night\_skiing\_state\_ratio, seems the most correlated with ticket price, this would mean seizing a greater share of night skiing capacity is positive for the price a resort can charge.

Visitors seem to put more value in guaranteed snow cover rather than more variable terrain area. Also, vertical drop seems to be a selling point that raises ticket prices as well.

Preprocessing and Model Training

After performing the preliminary assessments, we focused on creating a machine learning model, which was trained using a 70/30 train/test split. The scaled and the model was trained using Linear Regression and Random Regressor, using 2 separate pipelines. Performances were assessed using cross-validation technique, with and without feature scaling. We have found that the random forest model had lower cross-validation mean absolute error and was prone to less variability.  
  
  
Modeling

Once our model was created and trained, we used this model to calculate the expected Big Mountain Ticket prices. When using this model, Big Mountain Resort modelled price is $95.87 where the actual price is $81.00, this is with the expected mean absolute error of $10.39, which suggests there is room for a price increase. Please note, this model lies in the assumption that other resorts accurately set their prices according to what the market supports.   
  
To look at different scenarios for either cutting costs or increasing revenue, we looked 8 features that we found in our modeling:

|  |  |
| --- | --- |
| Vertical Drop  Graph shows that Big Mountain is doing well for vertical  drops, there are still a few resorts with a greater drop | Snow making area    Big Mountain is very high up the league table when it comes to snow making area |
| Total Number of Chairs    Big Mountain has amongst the highest number of total chairs, there also do seem to be a few outliers with extremely high number of chairs | Fast quads    Most resorts have no fast quads. Big Mountain has 3, which puts it high up that league table. Here are some values much higher, but they are rare |
| Runs    Big Mountain compares well for the number of runs. There are some with more but not many | Longest Run    Big Mountain has one of the longest runs. Although it’s just over half the length of the longest, the longer ones are rare. |
| Trams    The vast majority of resorts, such as Big Mountain have no trams. | Skiable terrain area    Big Mountain is amongst the resort with the largest amount of skiable terrain |

We also reviewed a few potential scenarios for either cutting costs/increase revenue.

**A graph of a price

Description automatically generated with medium confidenceScenario 1** - Where we observed potentially closing between 1-10 runs and analyze the change in revenue. If we were to close 1 run, it would make no difference in revenue. Closing 2 and 3 would reduces support for ticket price and so revenue. If Big Mountain were to close 3 runs, can can close up to 5 as there is no further loss in ticket price. If we were to close down to 6 runs, it would further reduce the ticket price.

**Scenario 2** - We also created a scenario where we added a run, while increasing the vertical drop by 150 ft and also installing an additional chair lift. This scenario increases support for ticket price by $1.99, and over the season this could amount to potential earnings of $3,474,638.  
  
**Scenario 3 –** We used the same criteria as scenario 2, except added 2 acres of snow making. However this yielded no further gain in ticket price, but could be associated with additional snow making costs.   
  
Price Recommendations  
With our model applied, it concluded that Big Mountain Resort modelled price is $95.87, while our current ticket price is $81.00. With our Mean Average Error of $10.39, it still accounts for an increase of $4.48, resulting in potential earnings of over $7.8 million.   
  
Also considering Scenario 2 if we where to make a one-time investment towards running a chairlift and adding a run that would give the mountain more than 150 ft of vertical drop would support an increase of ticket price by $1.99, resulting in $3,474,639 in seasonal earnings.

# Conclusion

As mentioned earlier Big Mountain Ski is always dedicated and focused on providing an exceptional experience for skiers and riders of all skill levels. Our market analysis of the features the Big Mountain Resort has to offer, indicates that Big Mountain has the capacity to be competing in the premium market segment. Our best model, the Random Forest Regressor, that was demonstrated can be easily deployed as a web app on the intranet for Blue Mountain Business Analyst team to utilize.

Our Future scope of work should be to analyze historical data pertaining to revenue, facilities, tracking changes in facilities, along with operating costs of every aspect of the facilities. This would allow us further enhance the modelled data with operational changes in revenue. Additional to that, scheduling a stakeholders meeting would be greatly beneficial to discuss any previous Blue Mountain price changes, market pressures and other historical financial data.