Guided Capstone Project Report

Big Mountain Resort

Problem statement

Big Mountain Resort is a large ski and snowboarding resort in Montana. Their goal is to identify its most and least valuable facilities in order to adjust their ticket prices to match their value while reducing unnecessary costs. They hope to increase revenue for the next season by at least $1.54 million, which covers the operating cost of the new lift.

Data Wrangling

The original data file contains information from 330 ski resorts across the U.S.

The range of weekend and weekday prices for each state can be found in Figure 1. Some states had a wide range, indicating a large variety of prices. Some tended to have weekend prices match their weekday prices. These variations could be accounted for by the variety of terrain, size of the state, the climate of the area, or proximity to mountain ranges. Montana specifically had a small range of prices and matching prices for the weekend and weekday.

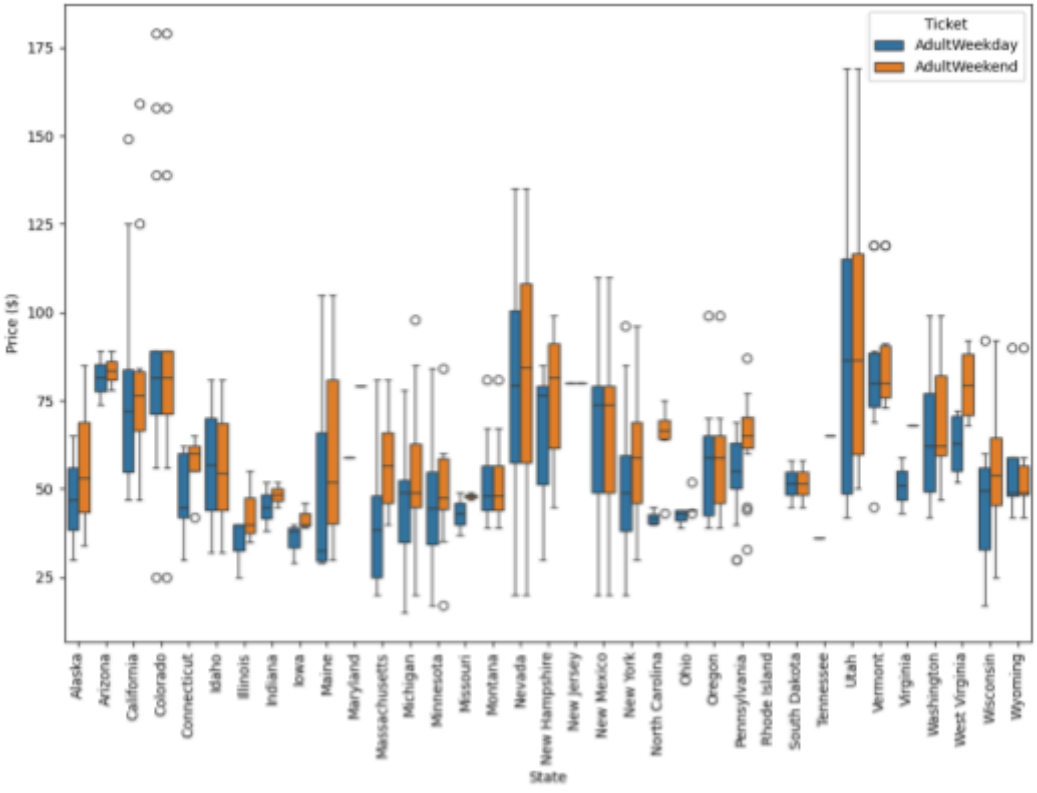


Figure 1.

A scatterplot of weekend vs. weekday prices show that a large number of resorts keep the same prices. But when they do differ, the weekend prices are higher than the weekday prices (Figure 2). Some resorts only provided the weekend price, and not the weekday price. The weekend price seemed to provide a more nuanced and complete picture of the data, so we used this value instead of the weekday price.

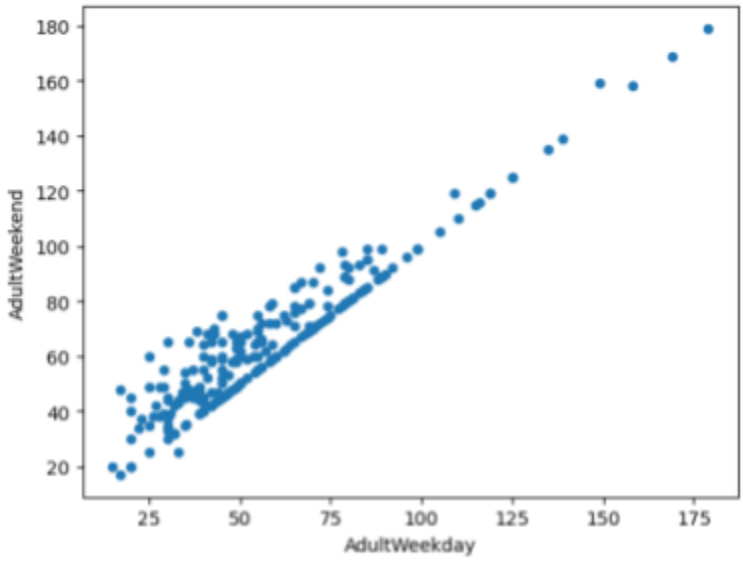


Figure 2.

We cleaned up the rest of the data by deleting missing and irrelevant columns, and sifting for and deleting incorrect values.

Exploratory data Analysis

Our goal is to determine which features positively impact the ticket price. We conducted a PCA (Principle Components Analysis) to look for patterns in correlation. We scaled the data, fit it into a PCA transformation, applied it to create derived features and explored more patterns. We calculated the ratios between values from individual resorts to the state-wide total, to get a better context of its relative significance, and created a new column based on the calculations.

The scatterplots below (Figure 3) helped us see which features were heavily correlated with a higher ticket price: the number of fast quads, the number of runs, the amount of snow made, how steep the vertical drop is, and the amount of nights skiing is available. We can conclude that customers want guaranteed snow, they want more access to the snow at night, and they want access to more trails (number of runs and fast lifts).

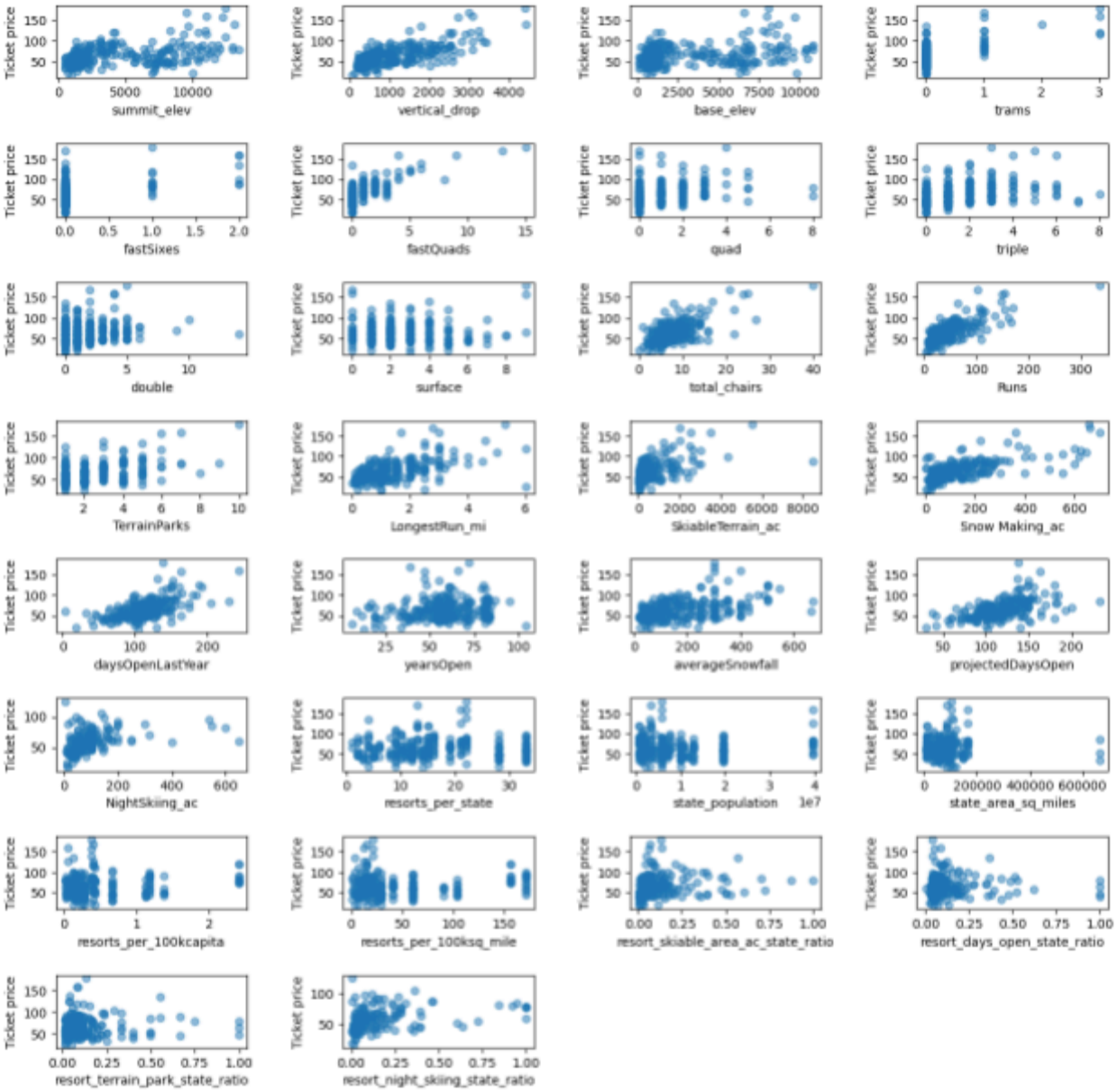


Figure 3.

Model Preprocessing with feature engineering

In the Preprocessing section, we split our data into 2 sets: 70% of the data went into training a model and the remaining 30% was reserved for testing the model. We found the mean absolute error based on the average, which told us that if we used the average as a baseline, then we can expect to be off by about $19 on average. Imputing missing values with the median and the mean improved the validity of the model. The mean absolute error decreased to $9 on average.

To refine the linear model, we focused on creating a nonlinear model based on the best k features. We used Pipeline to determine best features to make our model off of: the vertical drop

in feet, the amount of snow made, the total number of chairs, the number of fast quads, and the total number of runs in the resort park. (Figure 4)

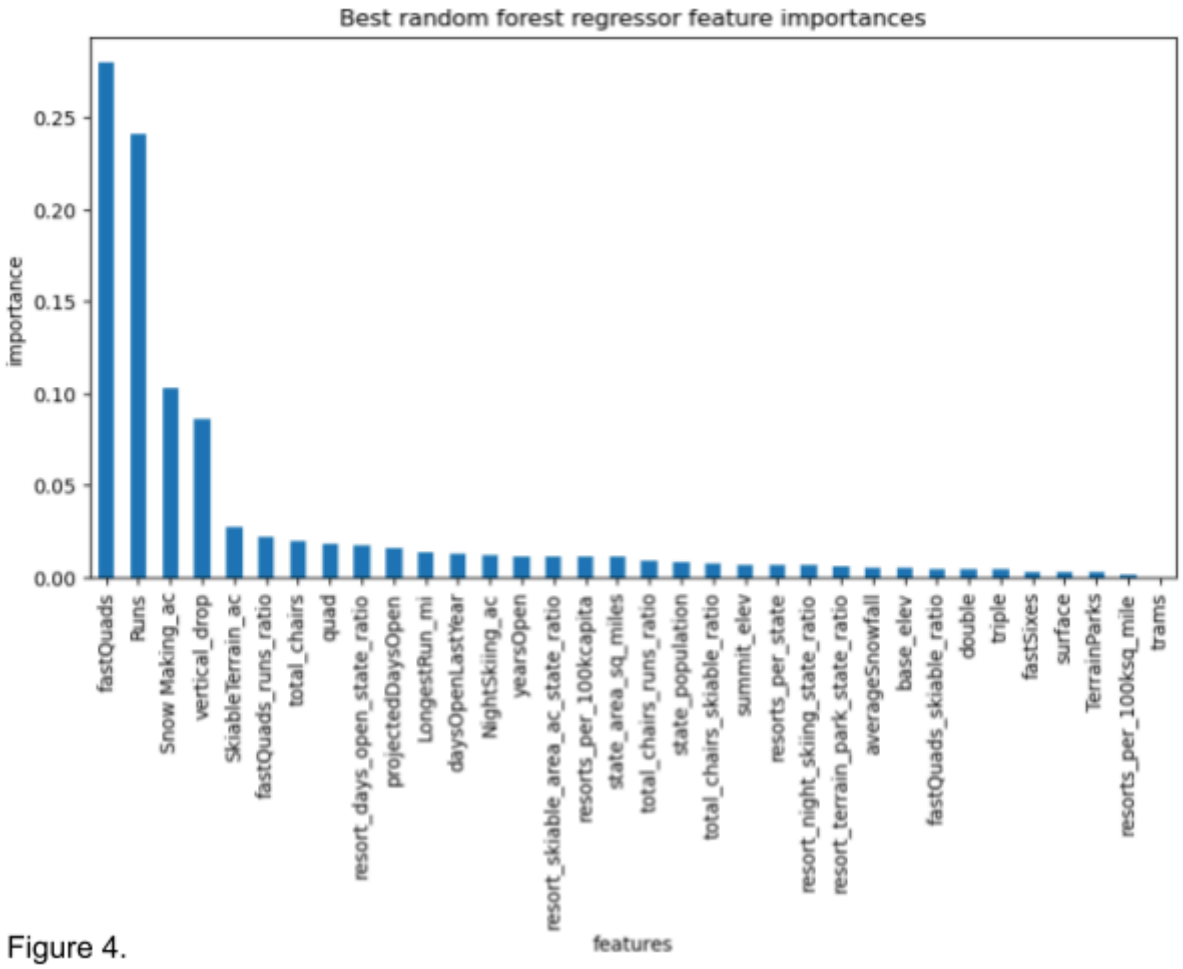


Figure 4.

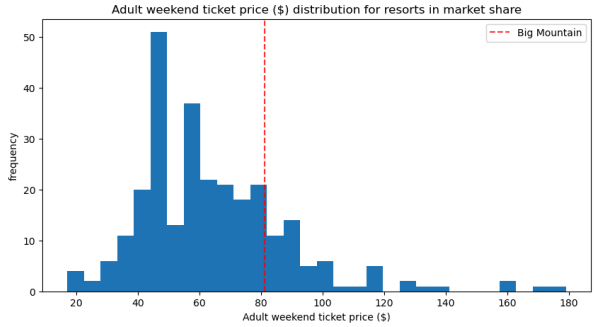
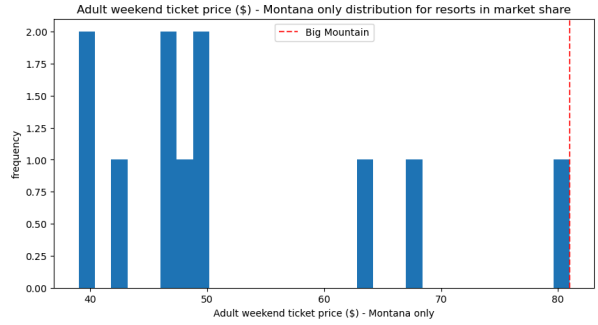
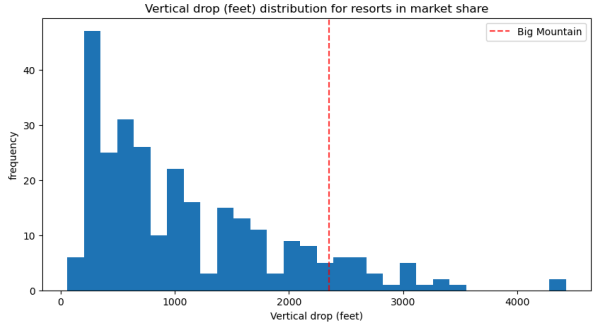
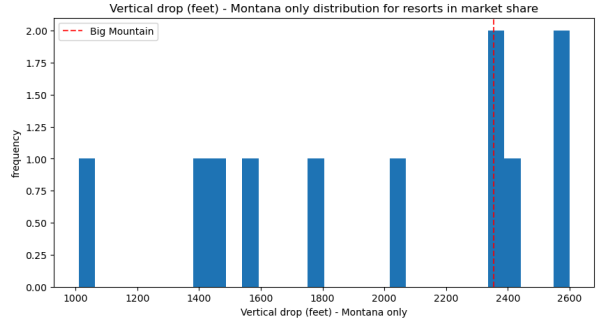
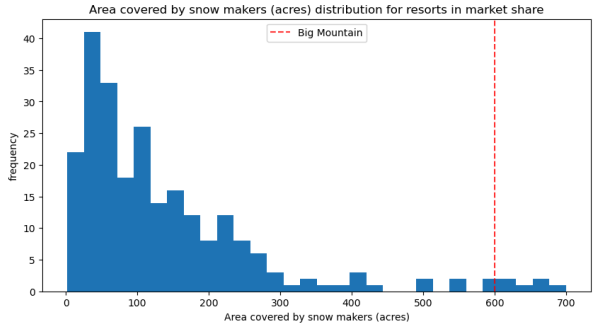
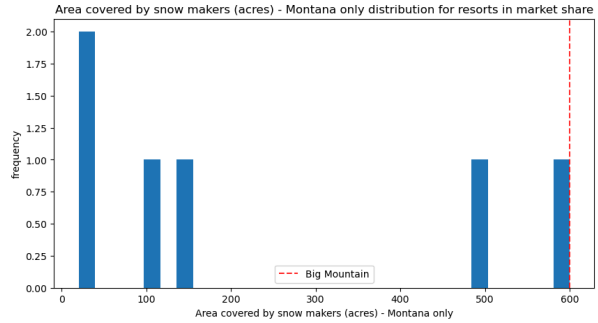
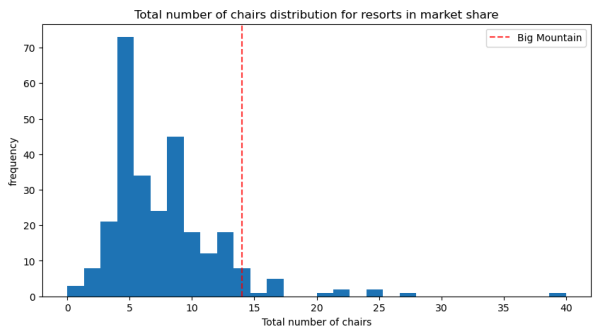
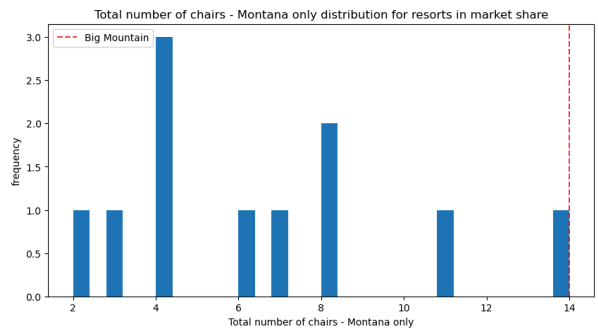
Algorithms used to build the model with evaluation metric

Both the linear and random forest model had consistent results, with slight variations. Calculating the mean absolute error for each model showed that the random forest model had less error (9.72) and the linear model (11.79). Therefore, we decided to use the random forest model.

Winning model and scenario modeling

Using a comparison graph and indicating where Big Mountain Resort lies relative to the distribution, it is clear that they reside in the upper percentile in the country for the features that matter: vertical drop, snow making area, total number of chairs, number of fast quads, number of runs, the length of their longest run, and the amount of skiable terrain they have.

Figure 5 also includes graphs of Big Mountain Resort's standing relative to its Montana competitors. It ranks #1 in all important aspects, (except in Total Number of Runs, in which it ties for #2 behind the Great Divide).

Nation-Wide Comparison	State-Wide Comparison
	
	
	
	

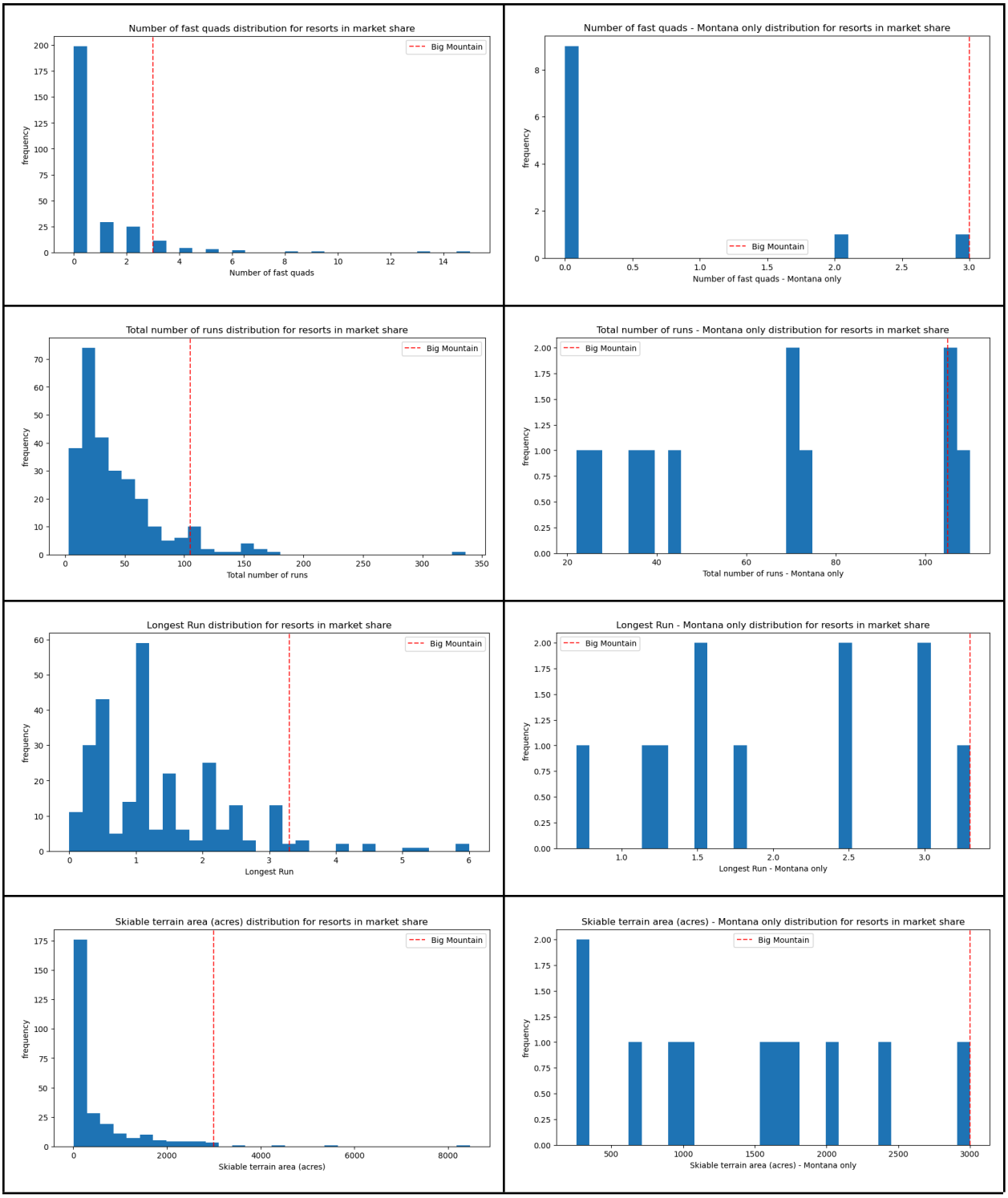


Figure 5.

Big Mountain Resort is expecting about 350,000 visitors a year. Assuming they buy on average 5 tickets a year, we can predict potential changes in revenue based on different scenarios.

If they close 1 run, they won't lose any money. If they close 2 or 3 runs, they will begin to lose more revenue. Closing 6 or more runs will have a dramatic impact on the revenue. Therefore, it is advisable to close one to five runs to save on costs. (Figure 5)

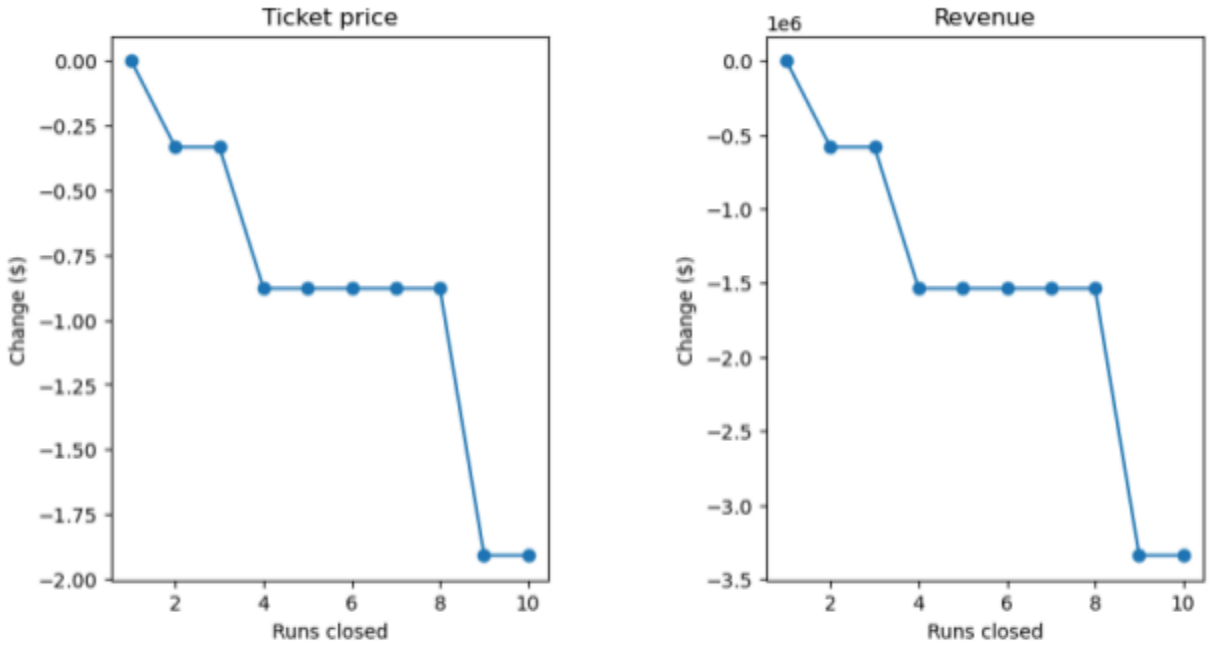


Figure 5.

The other possible changes (increasing the maximum vertical drop, adding 2 acres of snow, and increasing the longest run by 0.2 miles) don't seem to make a considerable impact on the revenue. Therefore, Big Mountain Resort should test out closing the least popular runs over time and see how that impacts the revenue.

Pricing recommendation

Currently, Big Mountain Resort charges $81 for their Adult Weekend price. However, according to the model and accounting for their relatively superior facilities, they could be pricing themselves at $97.85 (an increase of $16.85). The expected mean absolute error is $10.17, so even if they deviate from the model, they are well within their bounds of increasing their prices by $6.68 (for a total of $87.68).

Conclusion

Big Mountain Resort is one of the top ski resorts in the country and arguably the best in Montana. They consistently are in the top ranks for the most important features in a ski resort (reliability of snow, access to more of the mountain, amount of time open), and yet are charging the same amount as other local competitors. They have reason to increase their ticket price by $16.85.

Future scope of work

While this model has been helpful to gain new insights on what features are the most profitable and by how much, it could be improved with more information. Ideally, we would have the number of guests per year, to know how much final sales would be. This would also help us see

which resorts are more popular and have greater capacity. If possible, a breakdown of guests per month would also help us see which are peak months, and which months the operations can be shut down. Along this vein, it would be helpful to know how much the operations cost, including but not limited to: keeping certain trails open, the amount of electricity needed to operate lifts, the costs for making snow, and the salary costs for the employees and number of employees. If we were looking for long term solutions, I would also see how climate change impacted the resort's performance over the years and combine that with climate scientists' predictions to anticipate potential issues.