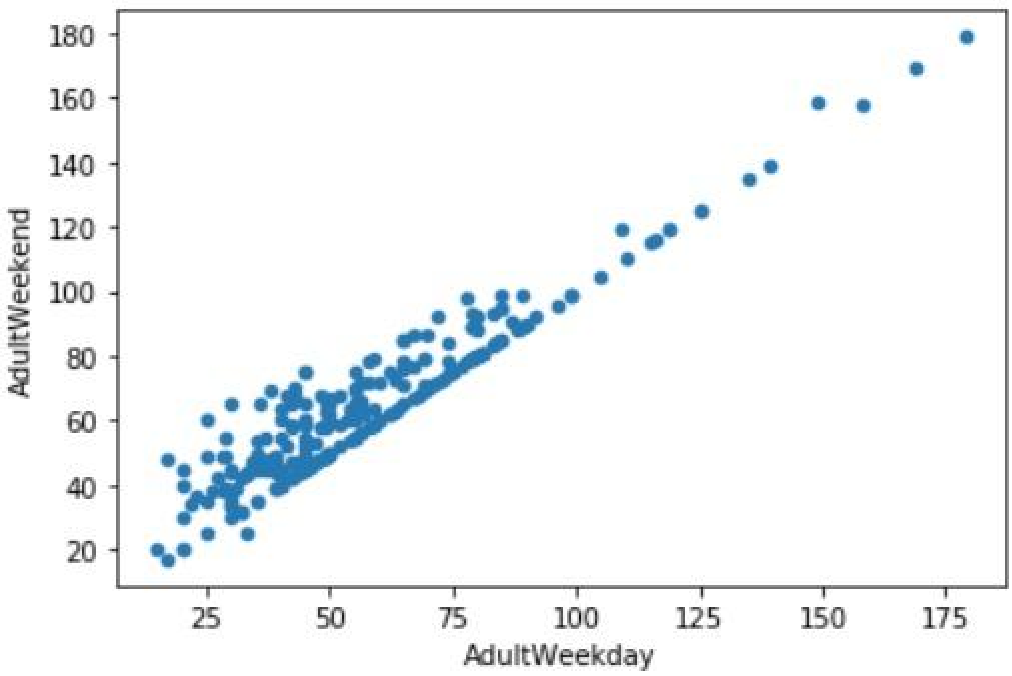
Big Mountain Resort project report
Springboard Data Science Track

Overview of big mountain resort:
Big Mountain Resort, a ski resort located in Montana. Big Mountain Resort offers spectacular views of Glacier National Park and Flathead National Forest, with access to 105 trails. Every year about 350,000 people ski or snowboard at Big Mountain.This mountain can accommodate skiers and riders of all levels and abilities.These are serviced by 11 lifts, 2 T-bars, and 1 magic carpet for novice skiers. The longest run is named Hellfire and is 3.3 miles in length. The base elevation is 4,464 ft, and the summit is 6,817 ft with a vertical drop of 2,353 ft. Big Mountain Resort has recently installed an additional chair lift to help increase the distribution of visitors across the mountain. This additional chair increases their operating costs by $1,540,000 this season.

Problem:
Big Mountain Resort is in need of a new pricing strategy, one that is based on data gathered from numerous ski resorts across the country. How can we create a pricing model that can determine a price that is competitive to customers, but also reflects the advantages of the facilities of the big mountain resort?

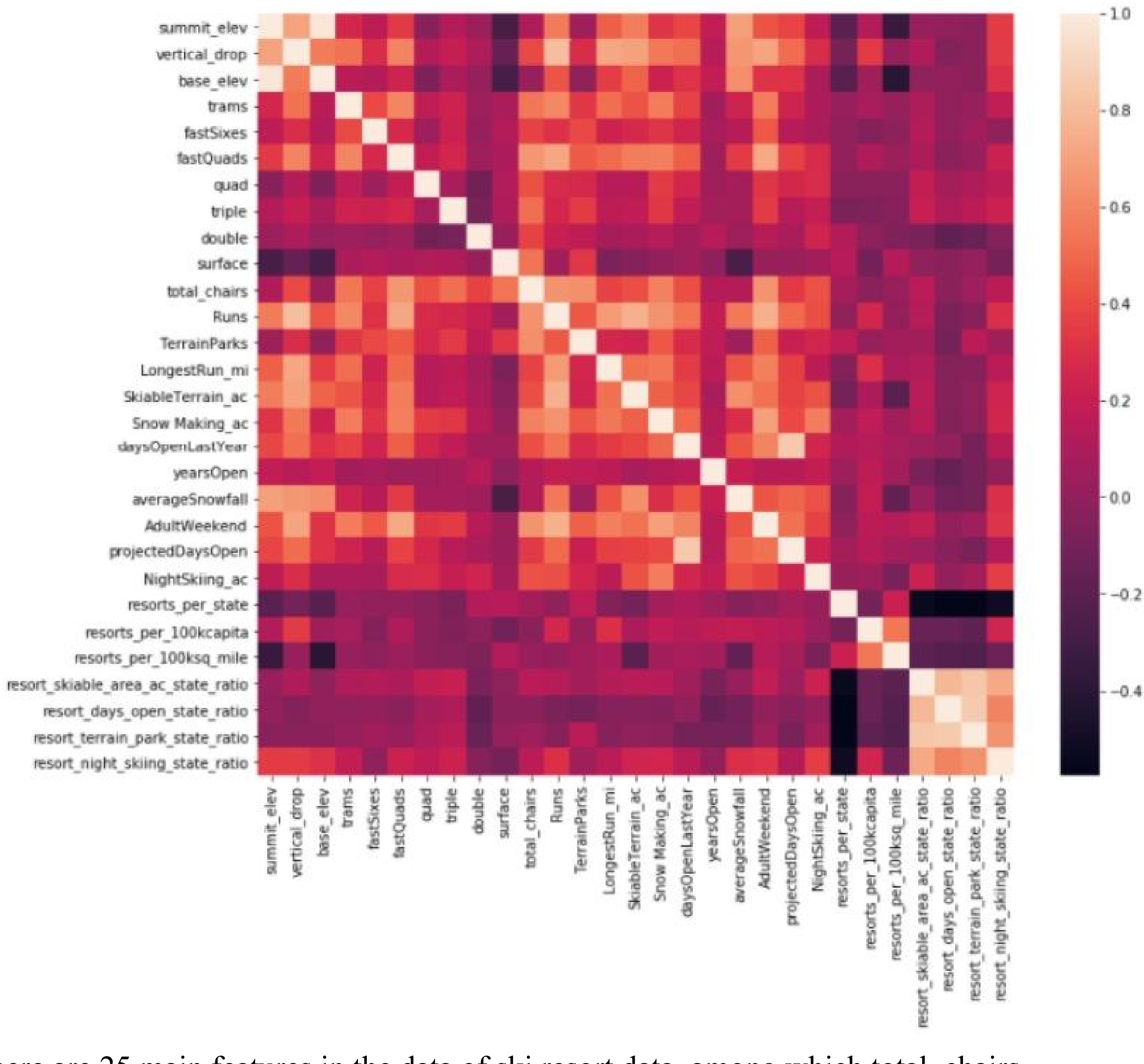
Data Wrangling:
Check the AdultWeekend vs. AdultWeekday prices to see how they are related and use the graph to see if marketing differences can be made by pricing weekends differently:



From the chart, we can clearly see that the prices of AdultWeekend and AdultWeekday in most states do not fluctuate too much, that is to say, their prices are roughly the same. There are a very large number of missing values, the two columns removed are 'AdultWeekday' and 'fastEight' because they have a height of zero. By processing these we are left with 277 rows.

EDA:

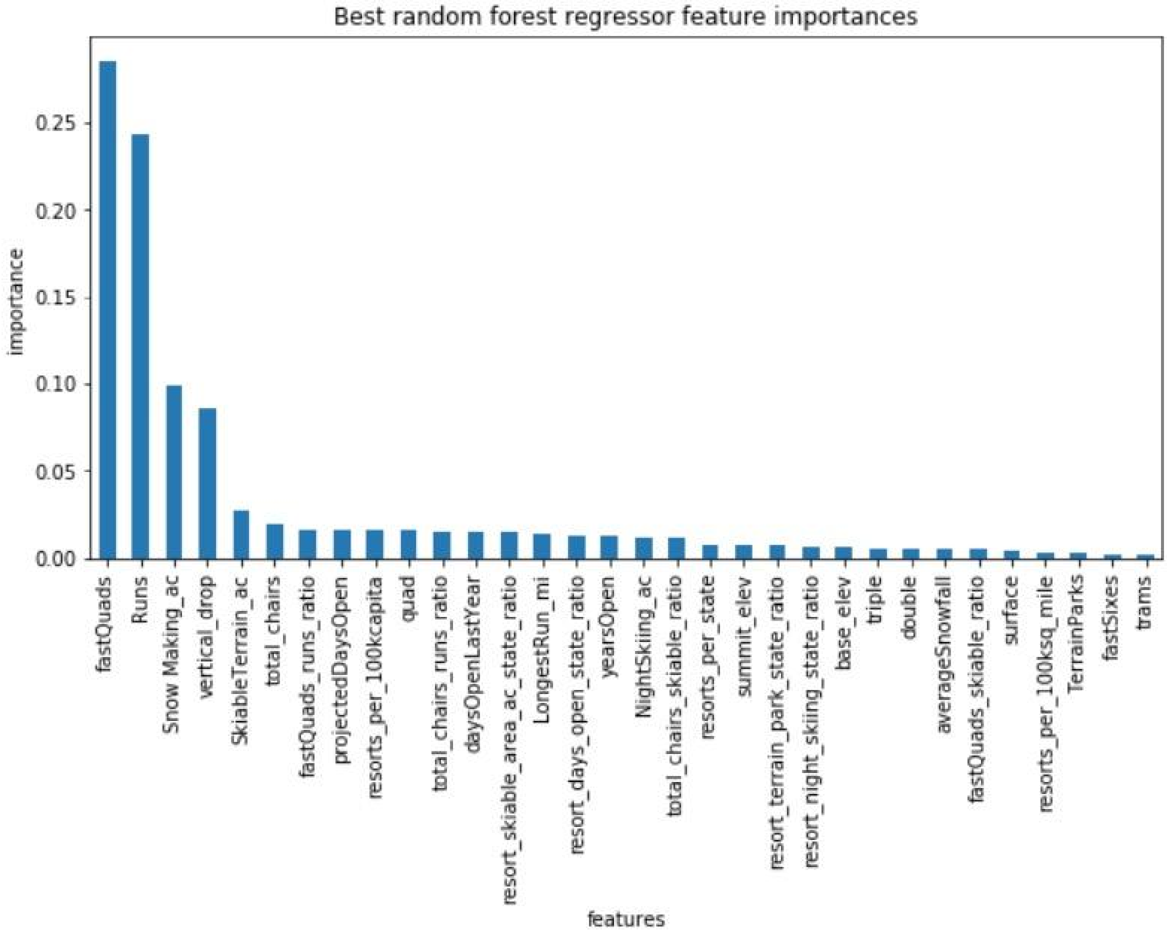
The heat map is the clearest relationship between price and facilities that can be seen intuitively. Turning attention to the target feature, AdultWeekend ticket price, we can see quite a few reasonable correlations. fastQuads stands out, along with Runs and Snow Making_ac. Visitors would seem to value more guaranteed snow, which would cost in terms of snow making equipment, which would drive prices and costs up.



There are 25 main features in the data of ski resort data, among which total_chairs, veritcal_drop, and runs are most relevant.Subsequent modeling needs to be considered mainly around the two main features of total_chairs and veritcal_drop. Improving these two data will directly affect ticket prices and operating costs. At the same time, the two features of total_chairs and veritcal_drop have a correlation with the fare.

Pre-Processing and Training Data:

Two modeling methods, linear regression and random forest regression, are established and compared. After comparing the data, we found that the random forest regression modeling method has a smaller error value, so random forest regression is a more suitable modeling method.



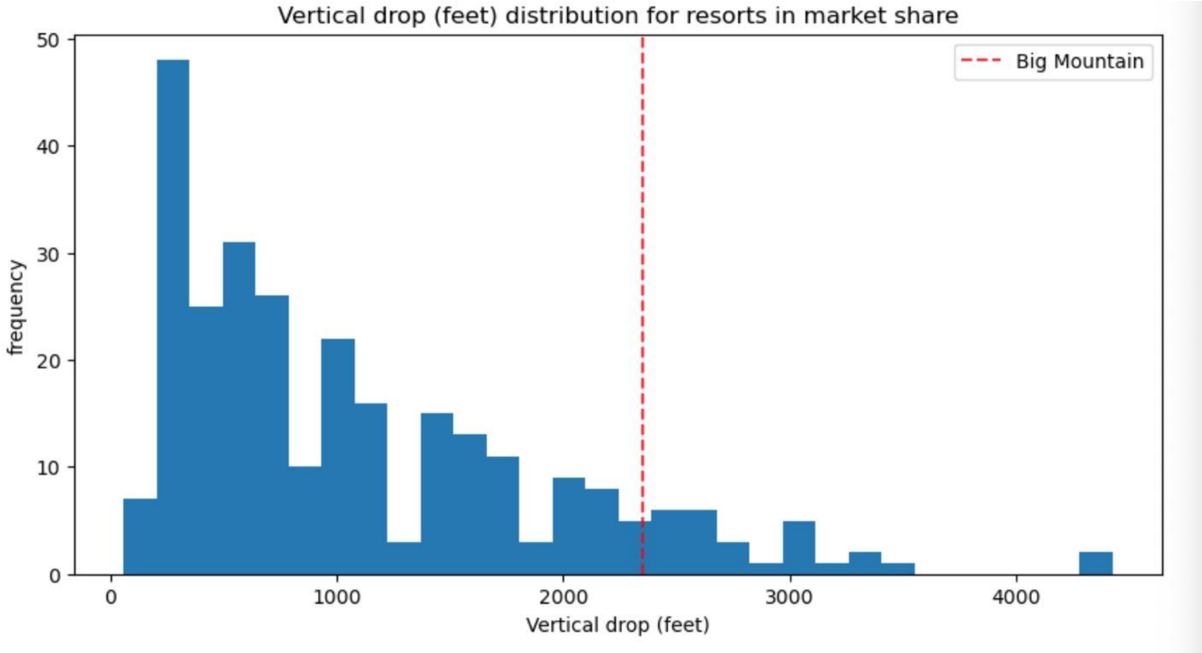
During the analysis we can see that the vertical drop also plays an important role in determining the fare. Adding this new component to the random forest model brought the mean absolute error down to around $1, the acceptable amount of variability. And use train_test_split to split the data.

Modeling:

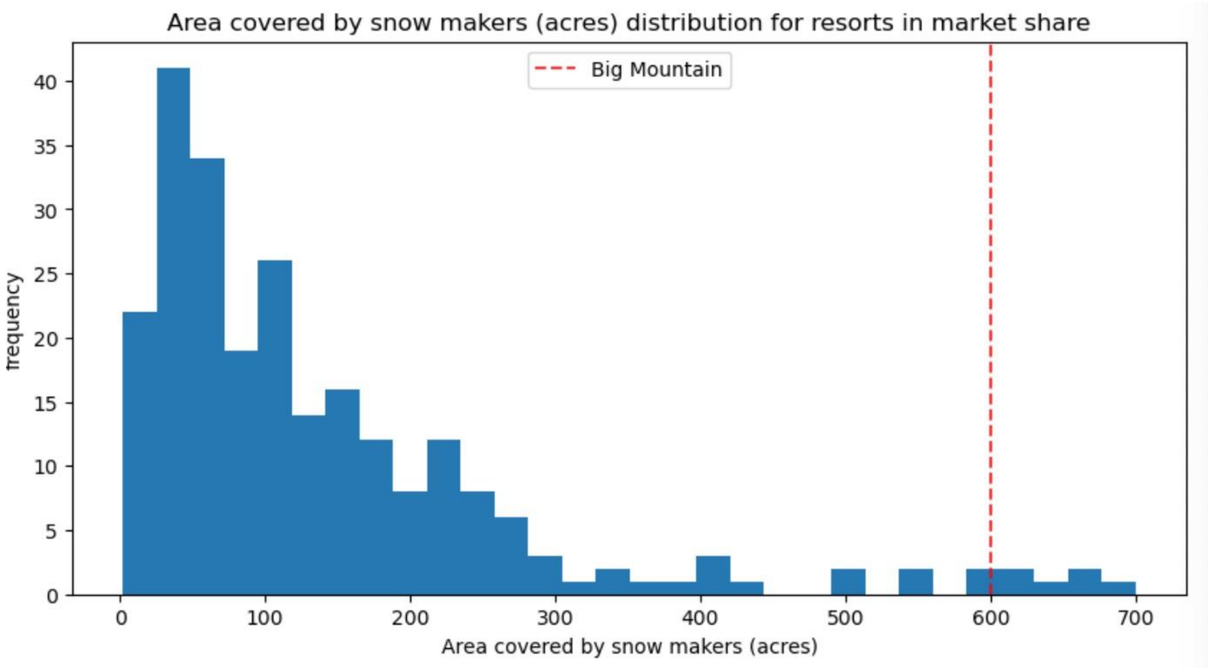
In the previous steps we have identified the regression method as well as the important completion components. To make the model more accurate, Features that came up as important in the modeling included:

vertical_drop
Snow Making_ac
total_chairs
fastQuads
run
Longest Run_mi
trams
SkiableTerrain_ac

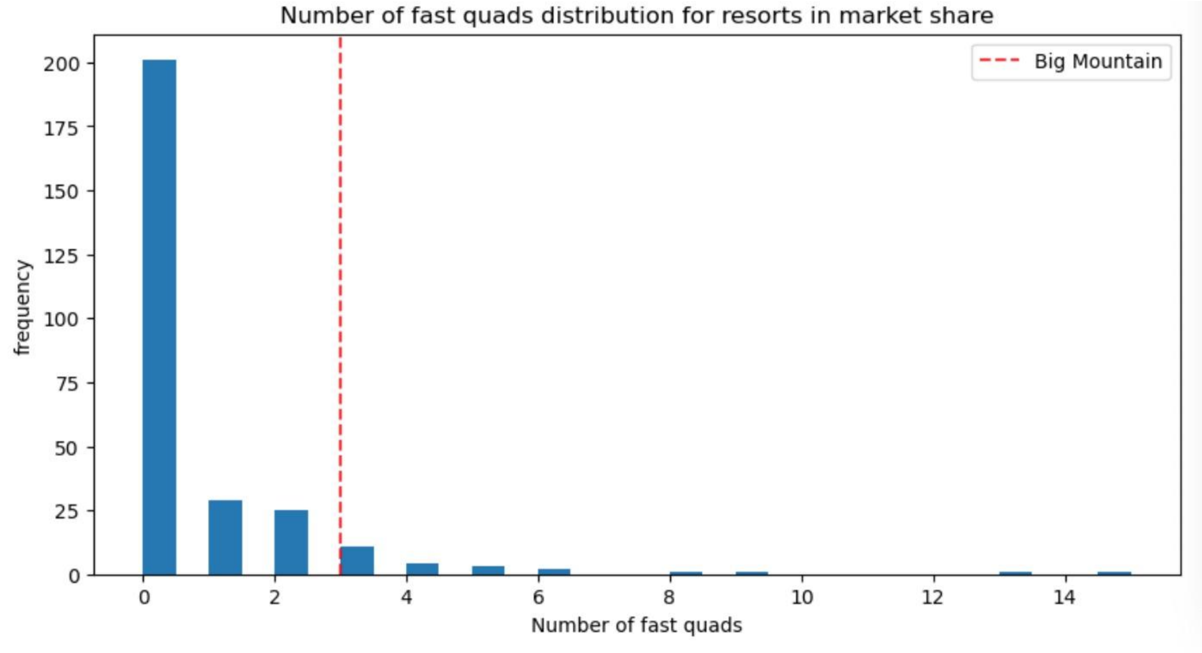
vertical drop:



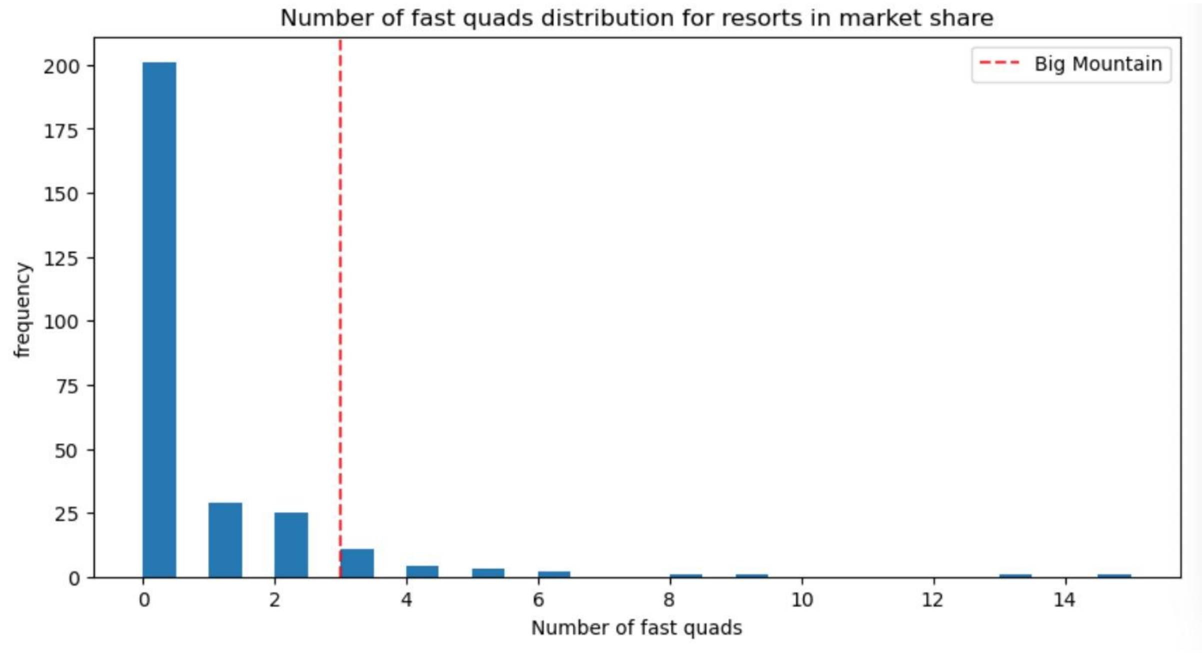
Area covered by snow makers:



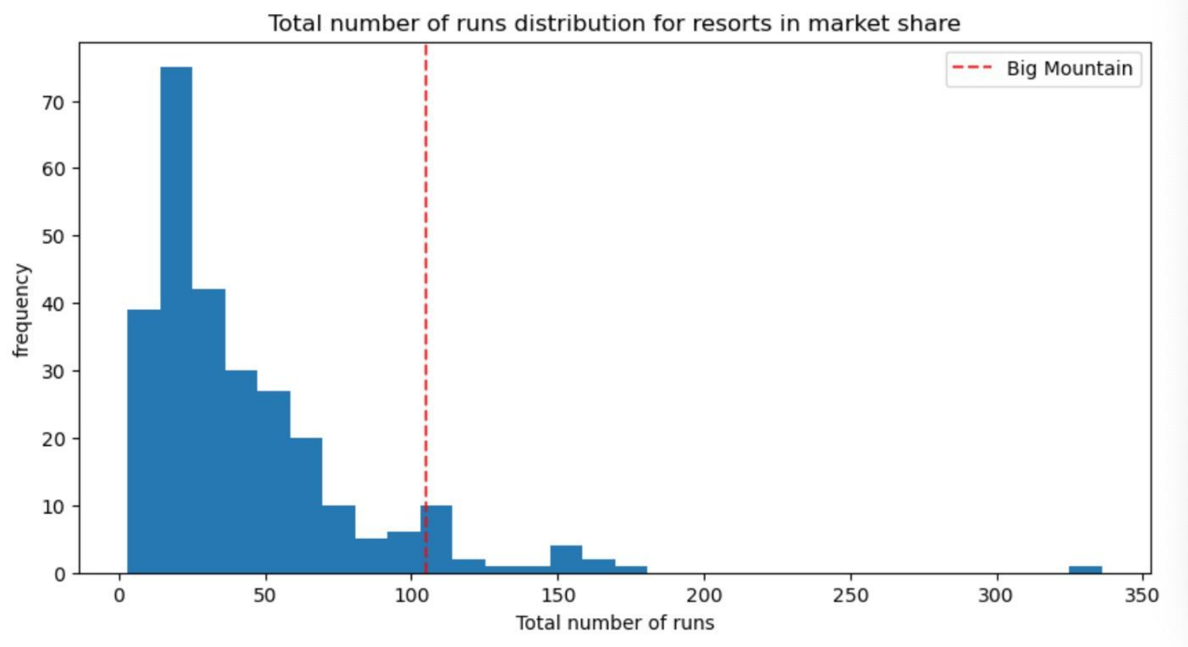
Total number of chairs:



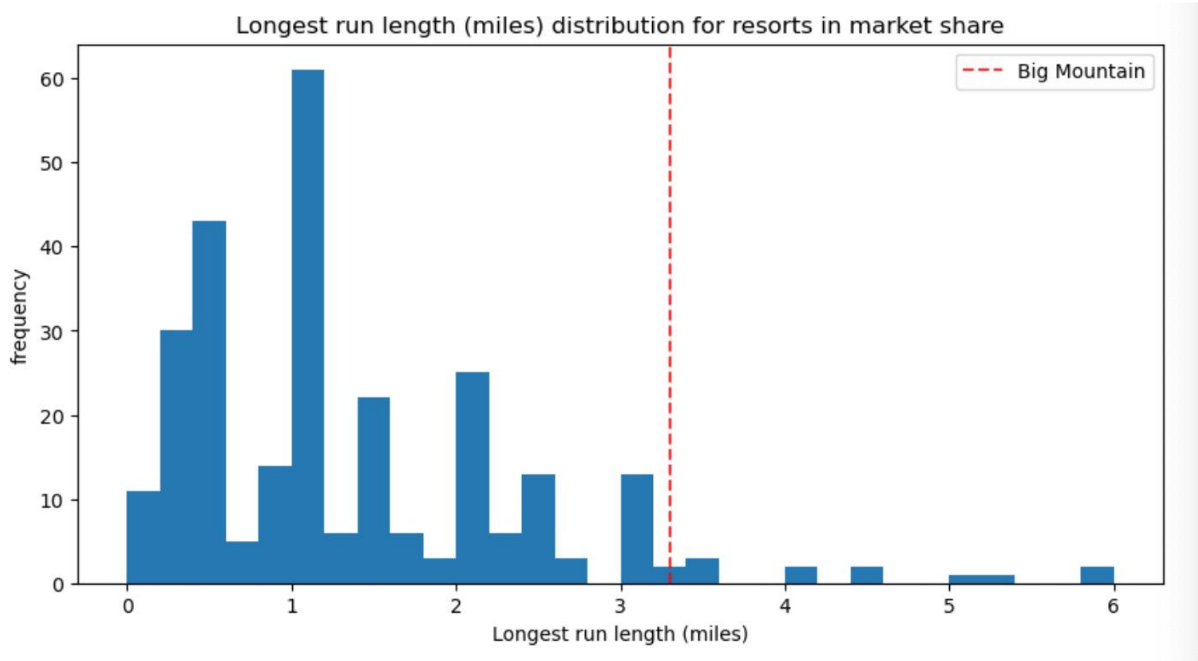
Number of fast quads:



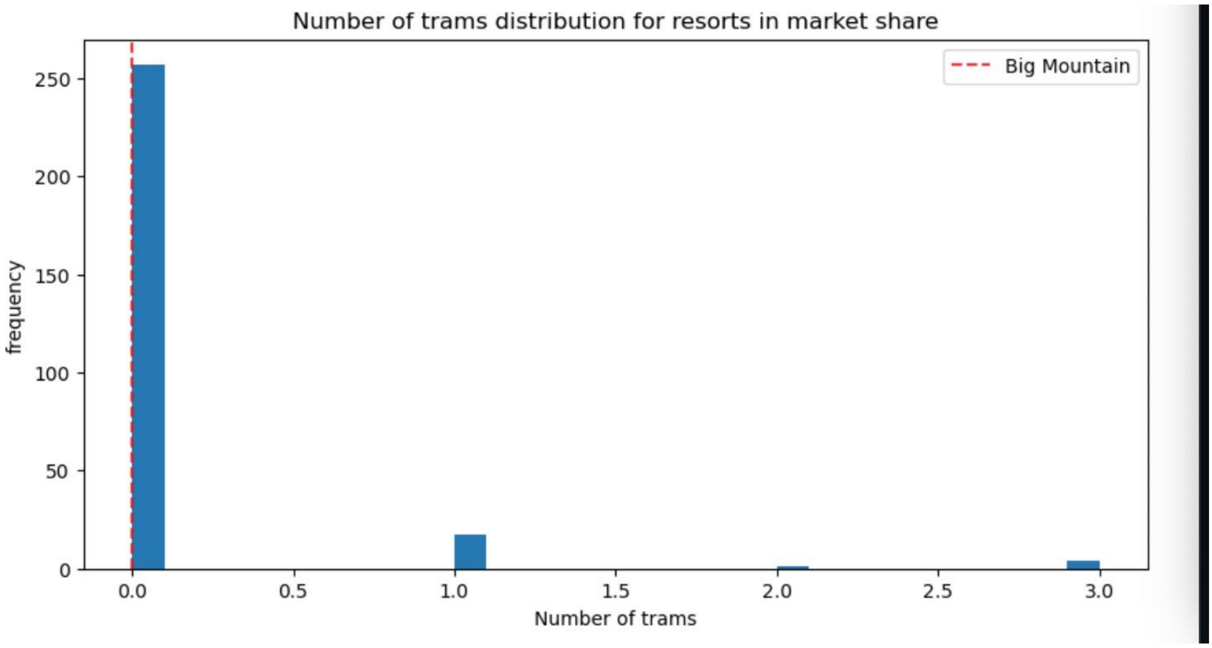
Total number of runs:



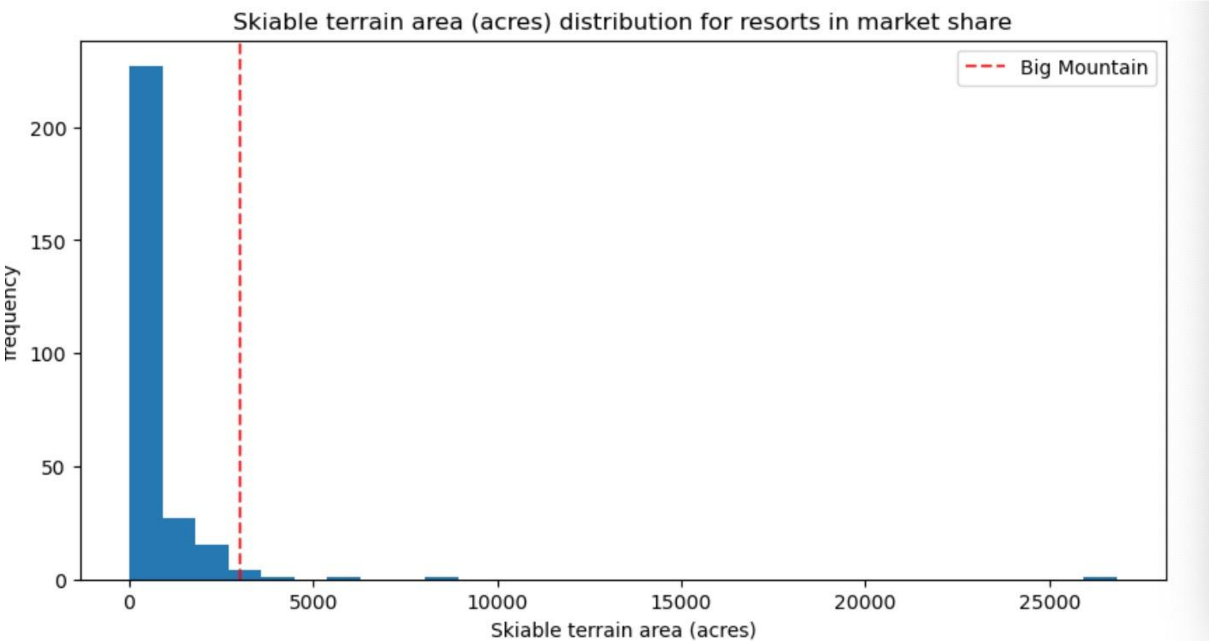
Longest run length:



Number of trams:

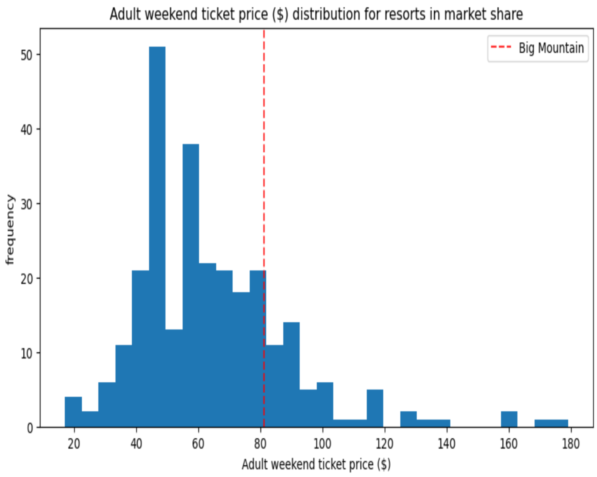


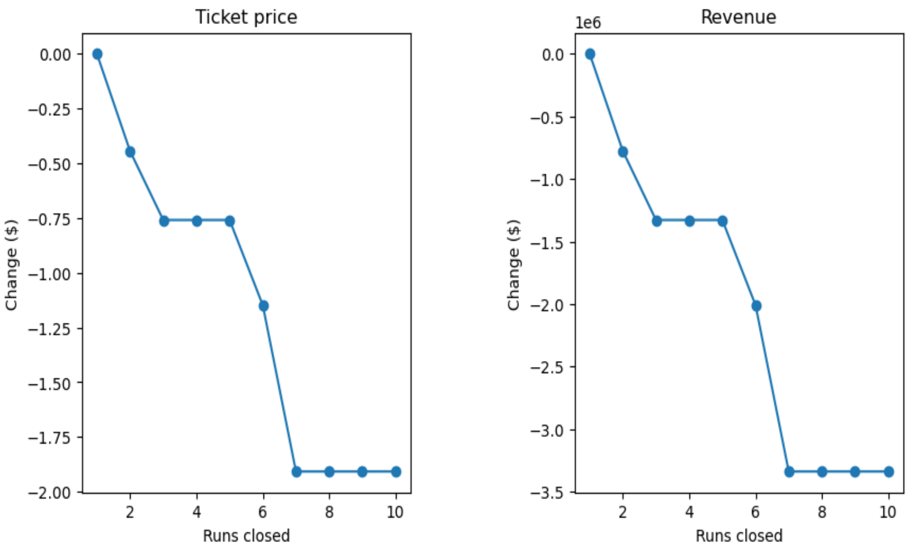
Skibale terrain area:



Through the previous case study, we can know that if we want to make Big Mountain Resort profitable, we must make two decisions by predicting a price increase/decrease or reducing operating costs. Big Mountain Resort modeled price is $92.96, actual price is $81.00. Even with the expected mean absolute error of $10.42, this suggests there is room for an increase.We can use this information to increase ticket prices to achieve profit (every $5 ticket price increase may increase the profit effect of $5-6 million). Or,increasing the vertical drop by 150 feet, and installing an additional chair lift，this scenario increases support for ticket price by $1.07. Over the season, this could be expected to amount to $1879630.Close up to 10 of the least used runs, up to five can make the price slightly lower, this is also one of the feasible release cases.

Conclusion:





By comparison, we should focus on the fare, because there is a difference of ten dollars between the standard fare calculated by the model and the actual fare. This is a very large gap, and ticket prices will be raised as the main release. It is reasonable to increase ticket prices by improving the quality of facilities. At the same time, the profit can be further increased by adding an additional chair lift. Reasonable model data results need to be handed over to company executives for evaluation. No matter what the executives think, the data is the most intuitive and convincing. Overall, I hope my analysis is accurate and effective, and I hope that big mountain resort can really be applied to business.