Big Mountain Ticket Pricing Model Report

Table of content

Problem statement

Big Mountain Resort (BMR) is installing an additional lift chair, which increases the operation cost to $1,540,000.00. BMR is searching for a pricing strategy to offset this increase in operation cost and to maximize its revenue while staying competitive with other skiing resorts in the market. BMR also looks for improvement plans in the future.

Data Wrangling

The ski resort data was provided by the Database manager, which includes the information about the features and pricings of the skiing resorts within the United States.
Two features are considered to be the target value, the ticket price during weekdays, AdultWeekday, and the ticket price during the weekend, AdultWeekend. A scatterplot between those two features indicates a strong 1-1 relationship between the two. Thus, one feature can be dropped out as the target features. AdultWeekday is dropped because it has more missing values.

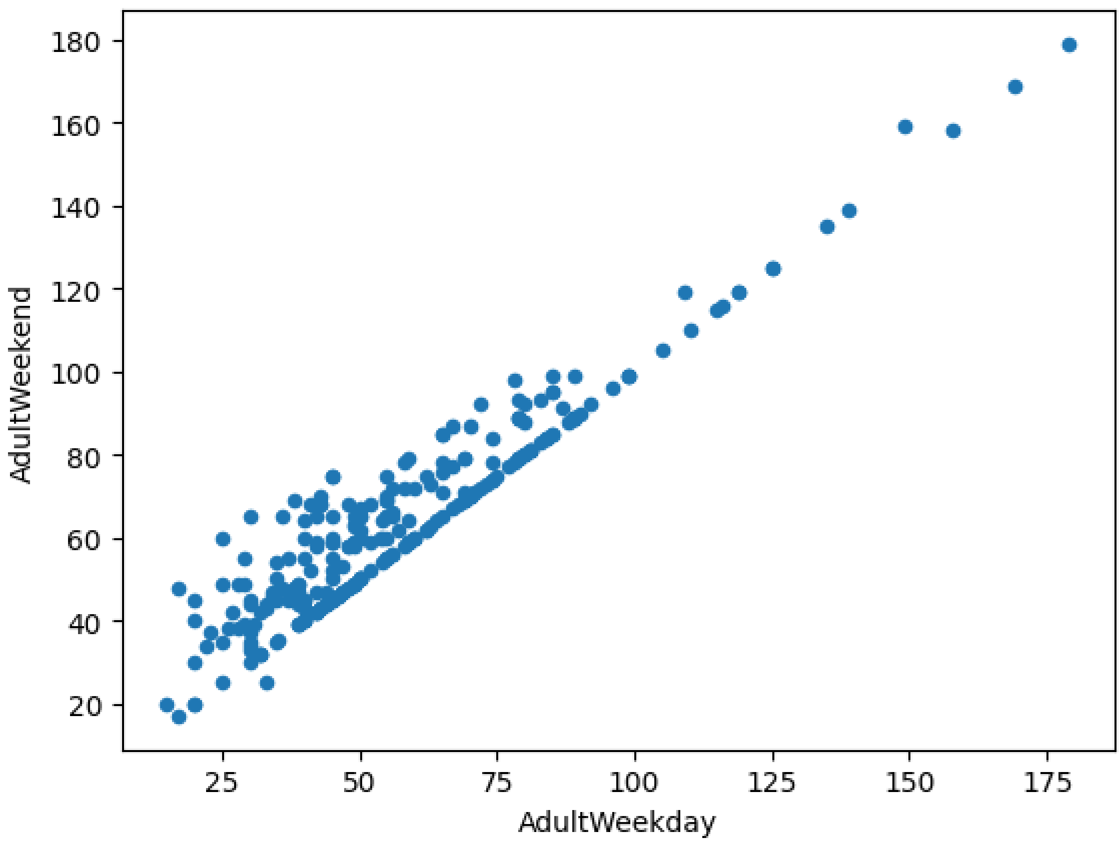


Figure 1: Scatter plot of AdultWeekday vs AdultWeekend

Rows with missing pricing data are dropped. Other data from the state, such as state population and state area, are added to the data set for further analysis.

Exploratory Data Analysis

One question remaining from the data wrangling process is to determine whether the model should be built on a nationwide scale or in a particular state. For this purpose, the state data is analyzed. Key features that are calculated from the state data are the Total state area, Total state population, Resorts per state, Total skiable area, Total night skiing area, Total days open and Resort density. Data is then scaled and applied PCA transformation. PCA transformation indicates that two features of state account for about 75% variance of the outcome.

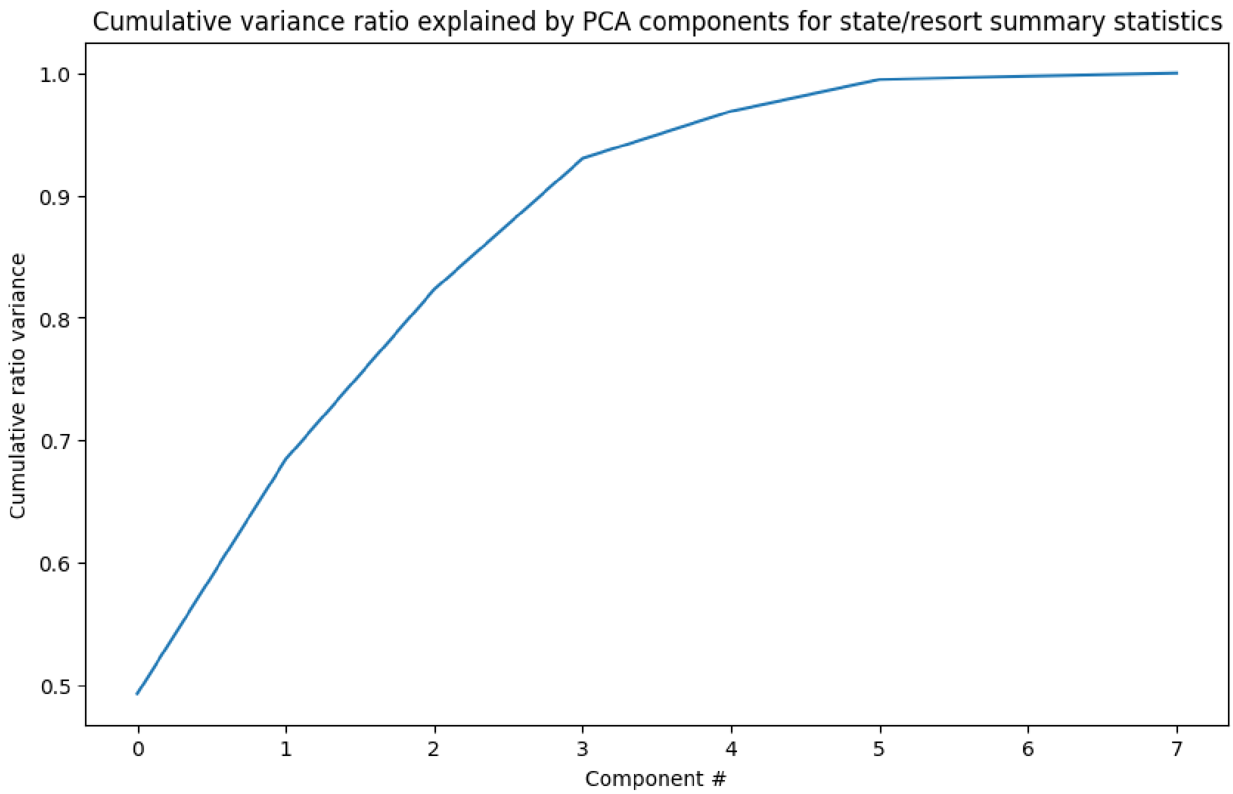


Figure 2: PCA Cumulative variance ratio of the state features

A histogram of average ticket price of each state is plotted to determine the distribution of the average ticket price. It seems that there is no special pattern for the state average ticket price.

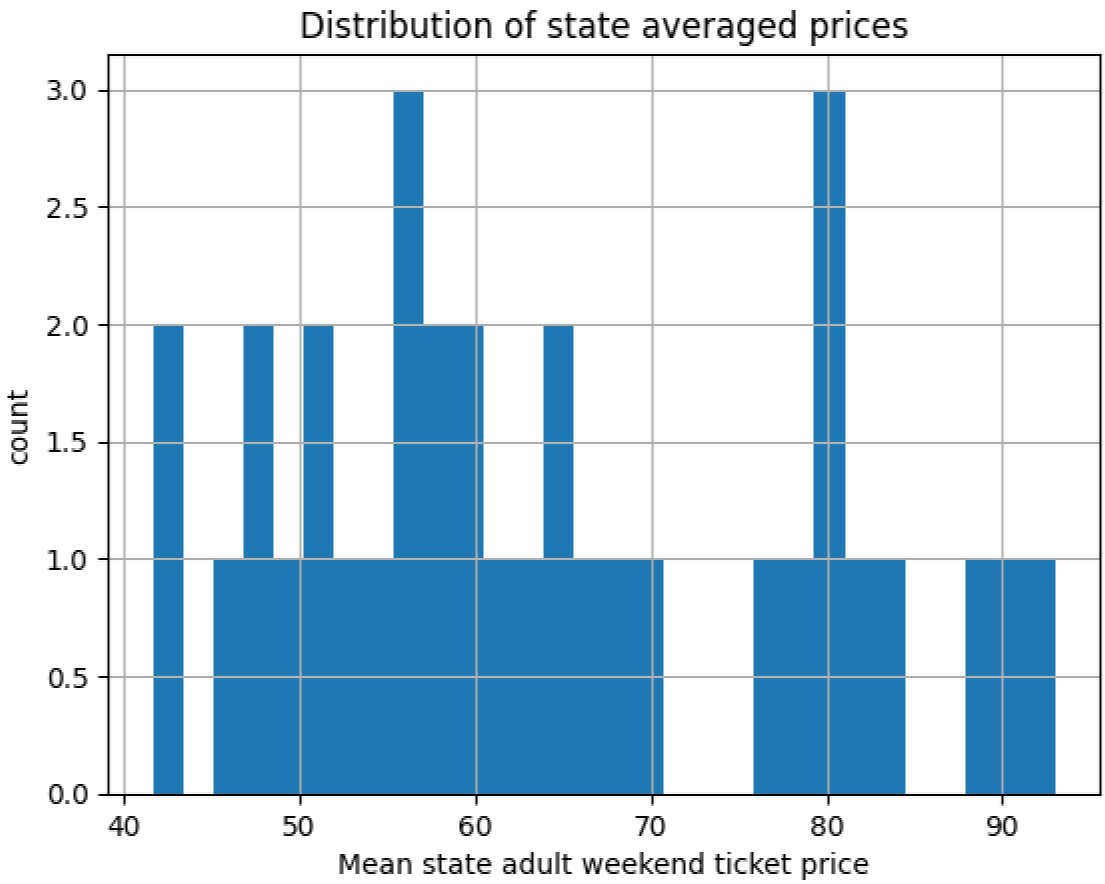


Figure 3: Average skiing resort ticket price in each state

Further analysis by adding state average ticket prices to the scatter plot of PCA first two components. The scatter plot does not indicate any special pattern.

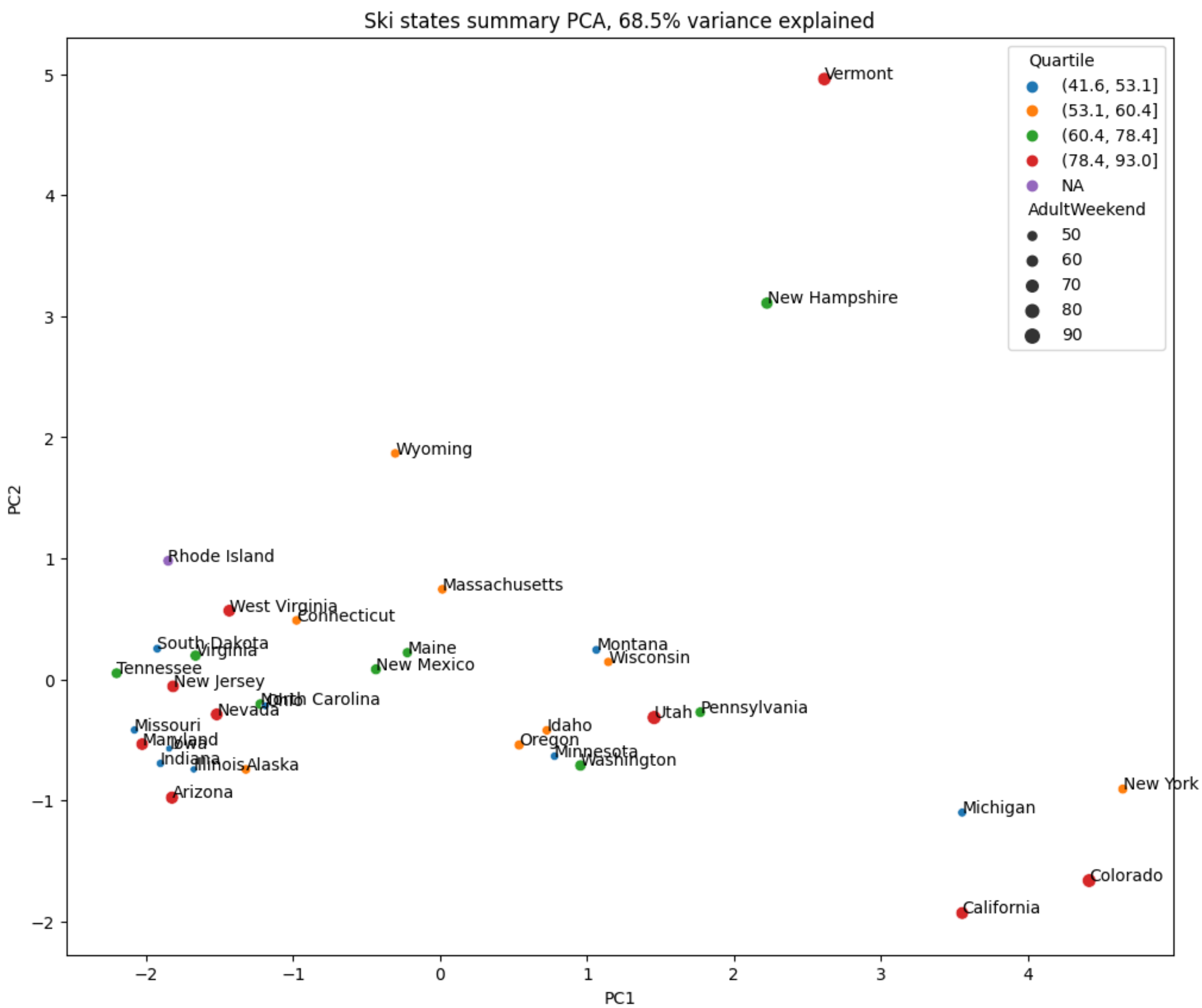


Figure 4: Average ticket price of skiing resort in each state

It can be concluded that the state label should be ignored for further analysis of the ski data set. Further analysis considers the correlation between numeric features in the ski data set with respect to the target value. Heatmap of features and scatterplot of each feature with respect to the target value is drawn.

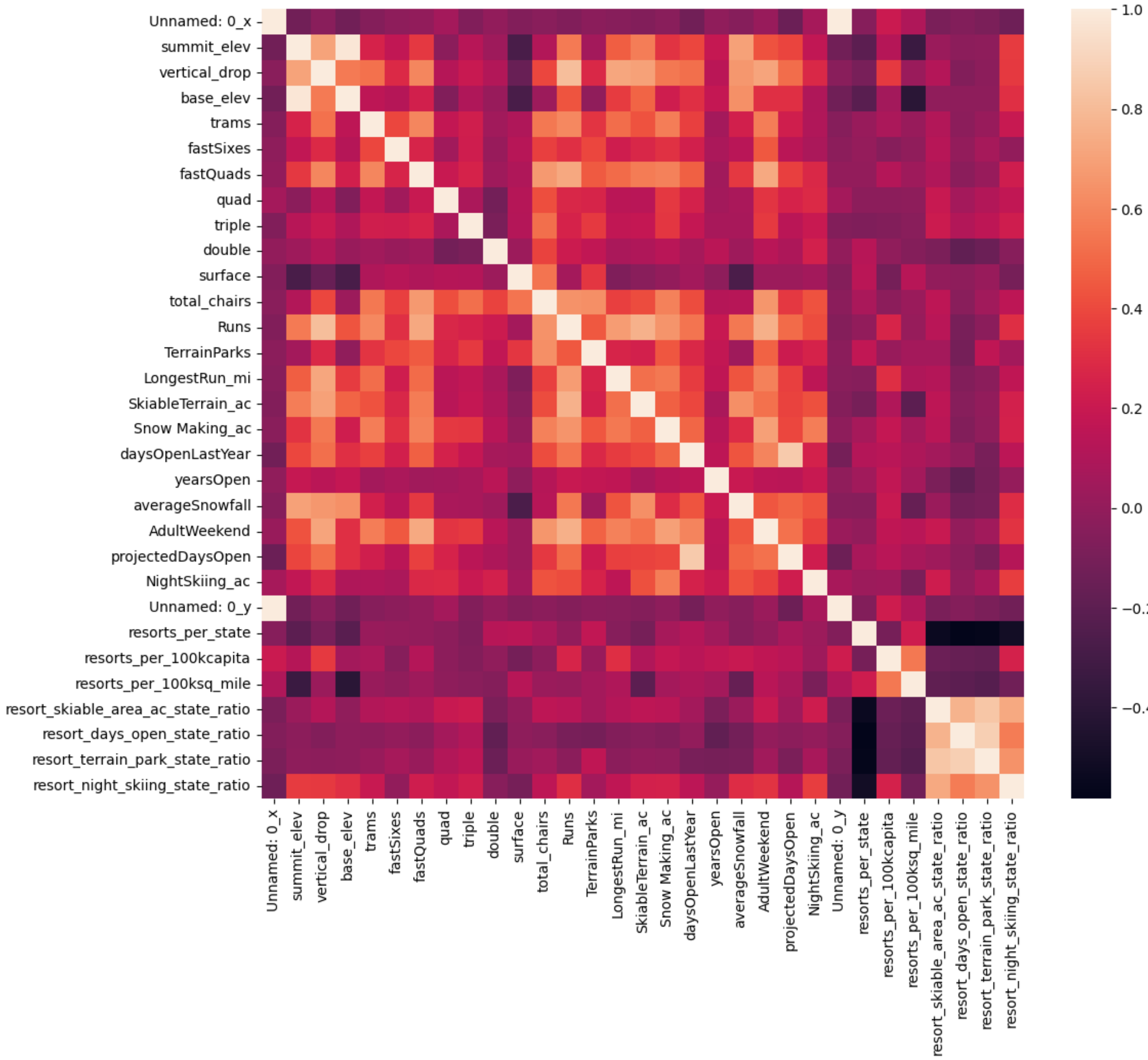


Figure 5: Heatmap of features in ski data set

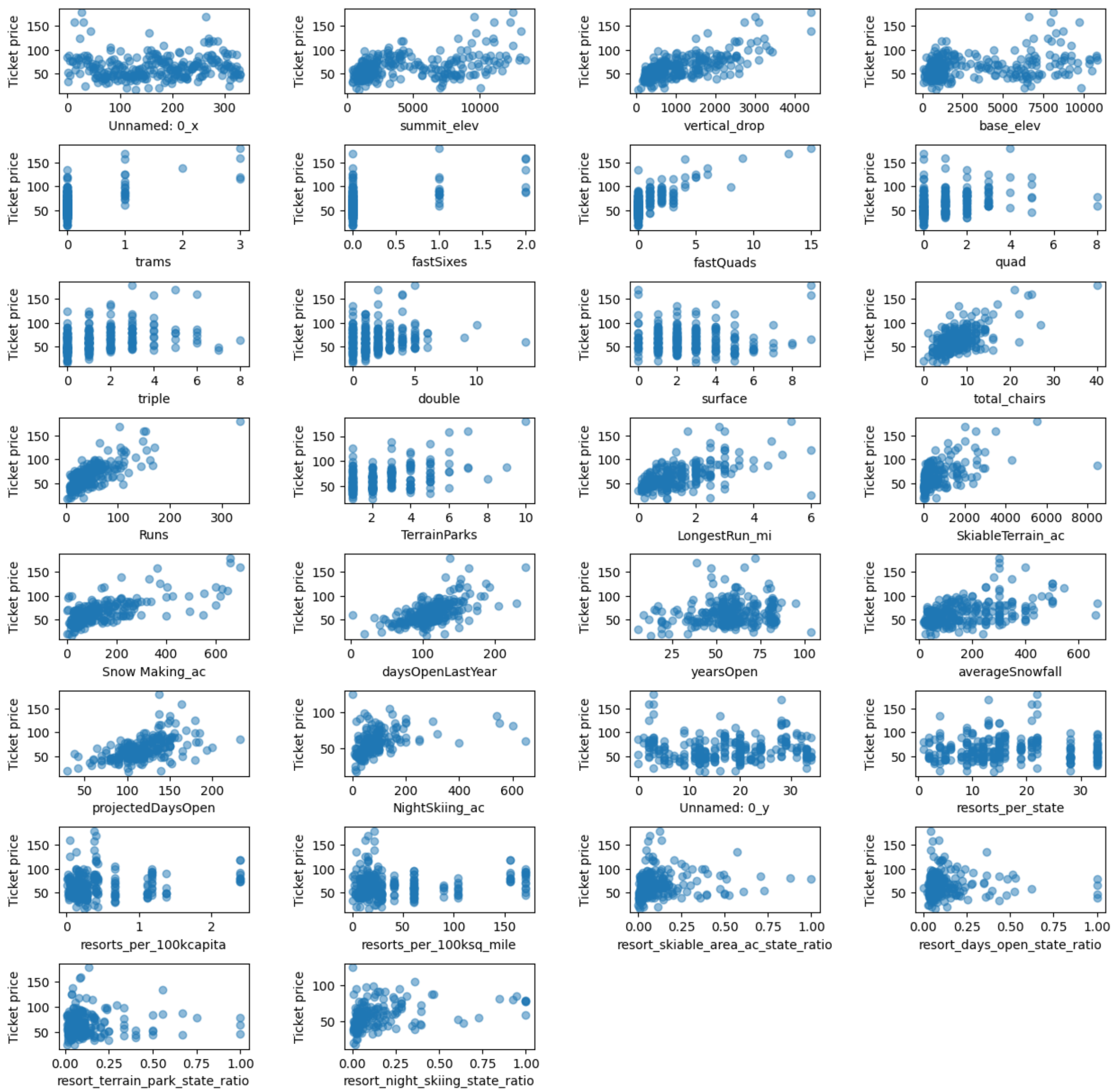


Figure 6: Scatter plots of each feature in ski data set with respect to ticket price

From the heatmap and scatter plots, four features show strong correlation with the target value are vertical_drop, fastQuads, Runs and total_chairs.

Model Preprocessing with feature engineering

The ski data set is then divided into test and training sets with a ratio of 70/30 for preprocessing and training. A based model is built using the mean value of the ticket prices as predicted values. Three metrics used to assess the models are R-square (coefficient of determination), Mean Absolute Error (MAE), and Mean Square Error (MSE).

Two strategies to impute missing values are mean and median. Pipelines are created to train data using linear regression and random forest regression. Cross-validation is also applied in the training process.

The linear model with cross validation indicates eight most important features in determining the target value, which are vertical_drop, Snow Making_ac, total_chairs, fastQuads, Runs, LongestRun_mi, trams, and SkiableTerrain_ac.

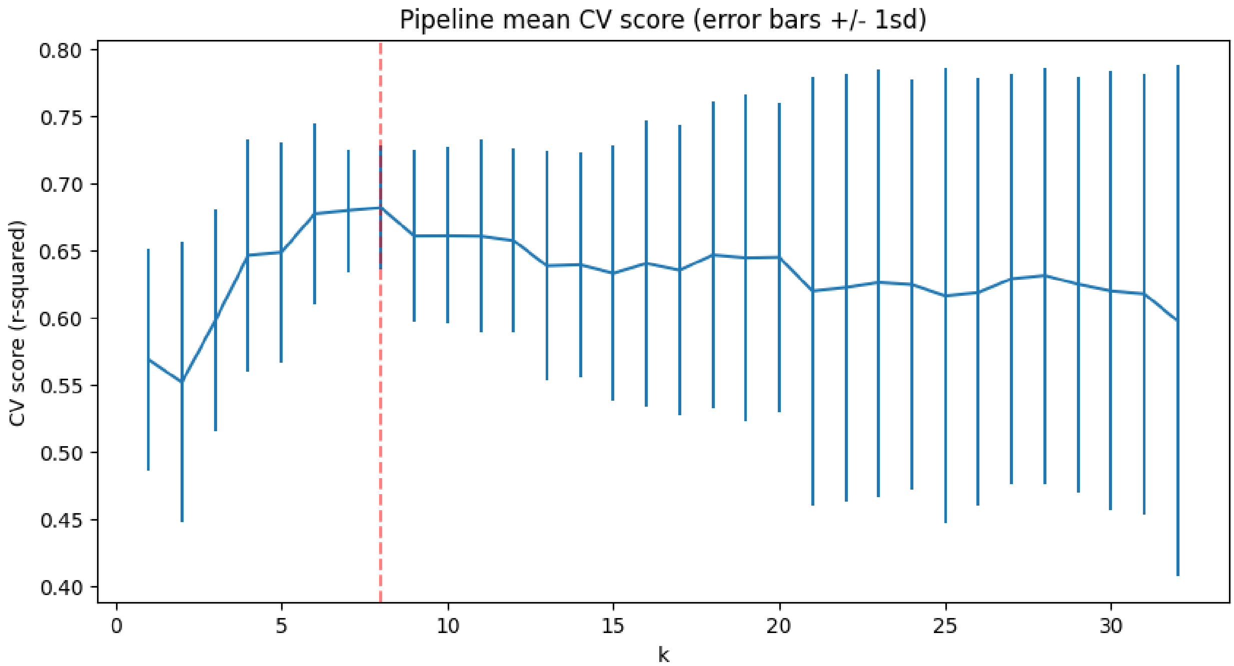


Figure 7: CV score of K component in the linear regression model

The random forest model indicates four most important features in determining the target values, which are fastQuads, Runs, Snow Making_ac, vertical_drop. These features agree with the linear model.

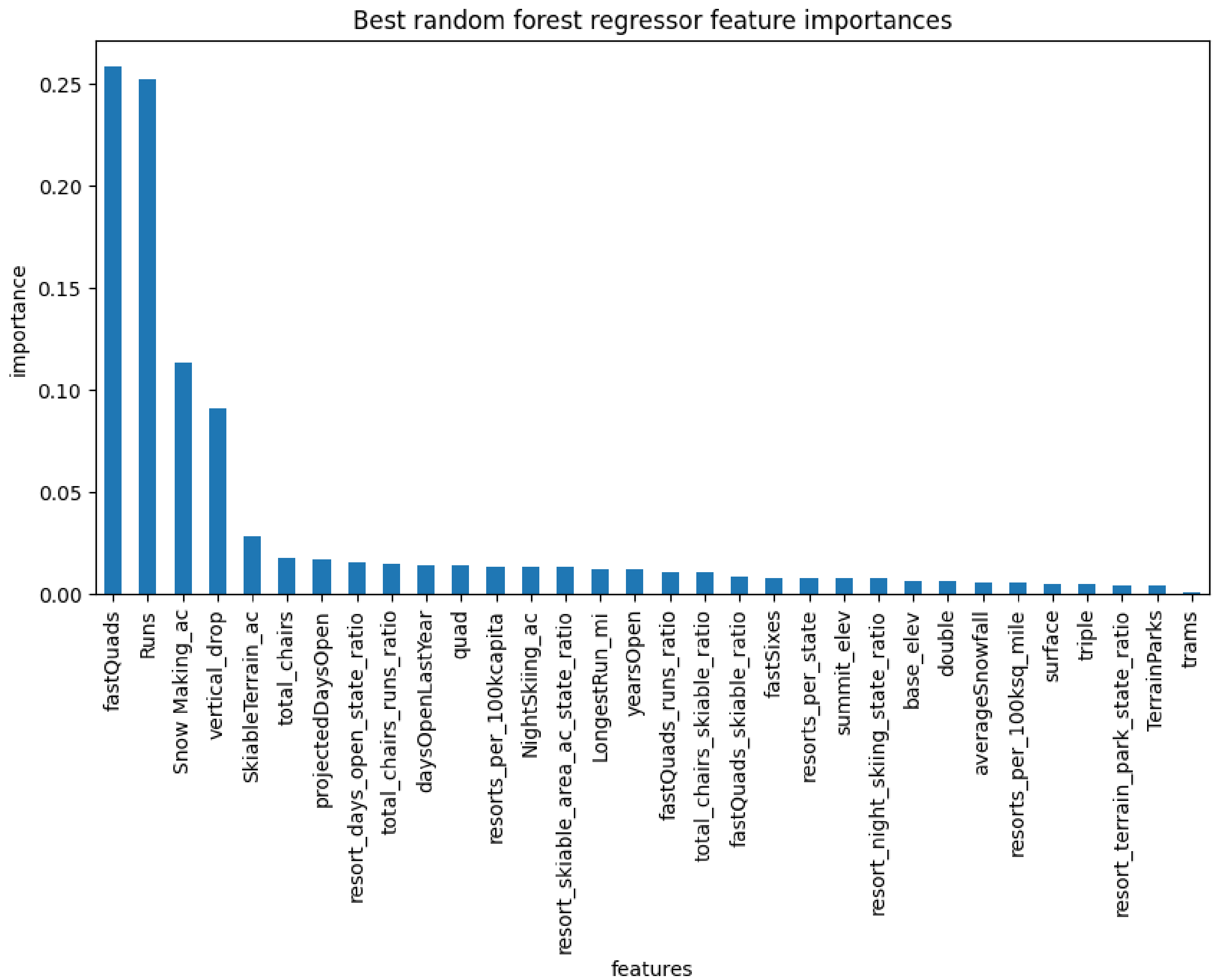


Figure 8: Importance value of each feature in the random forest model

From the linear regression models and random forest models with cross-validation, the best model of each method is chosen and compared to each other

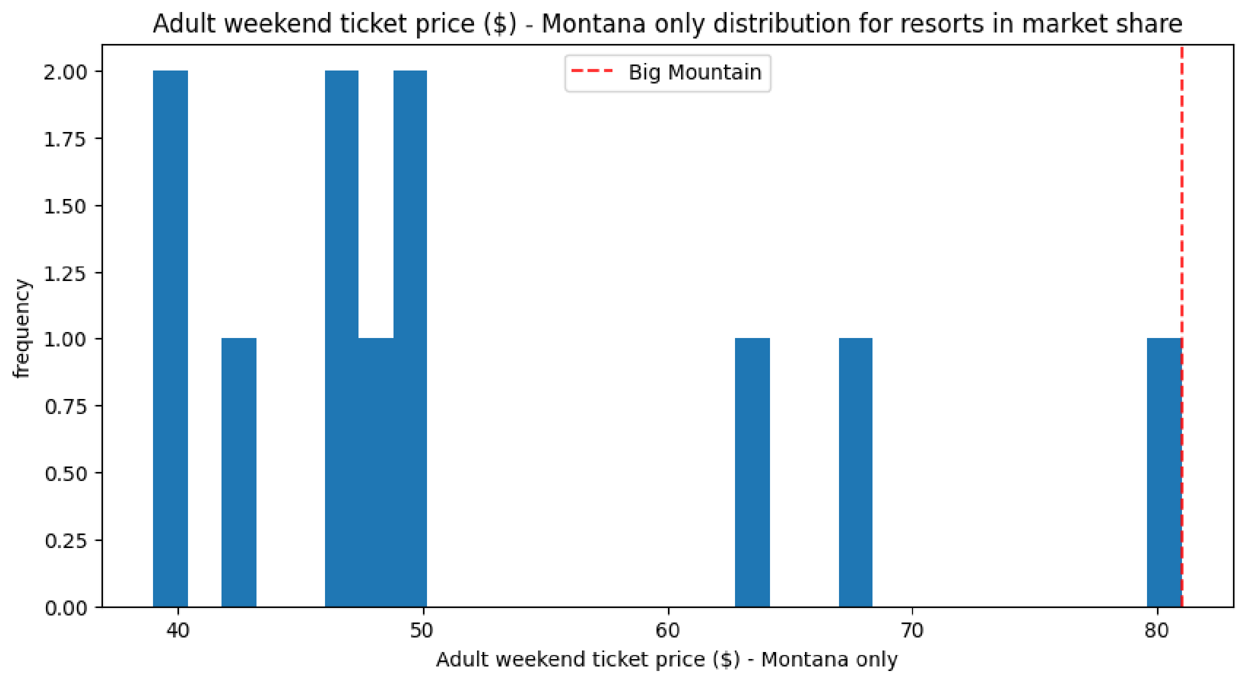
1. Linear regression model: Mean Absolute Error: 11.793465668669324
2. Random forest model: Mean Absolute Error: 9.537730050637332

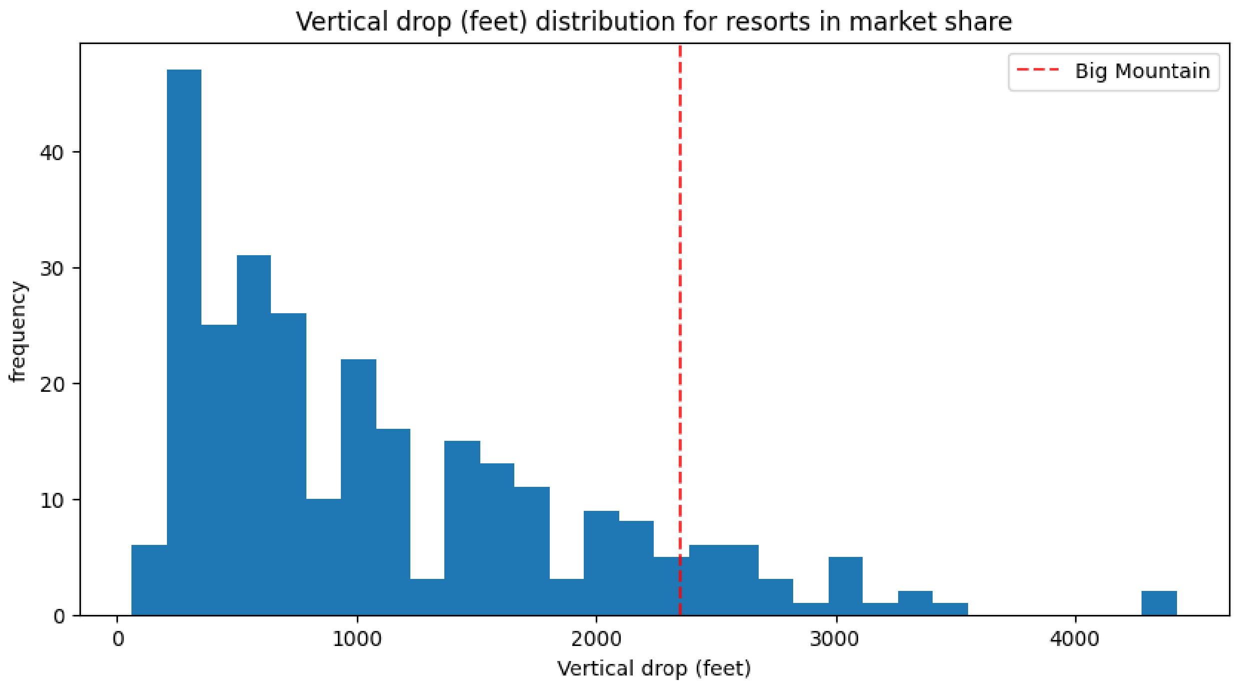
The best random forest model is chosen to predict the target value.

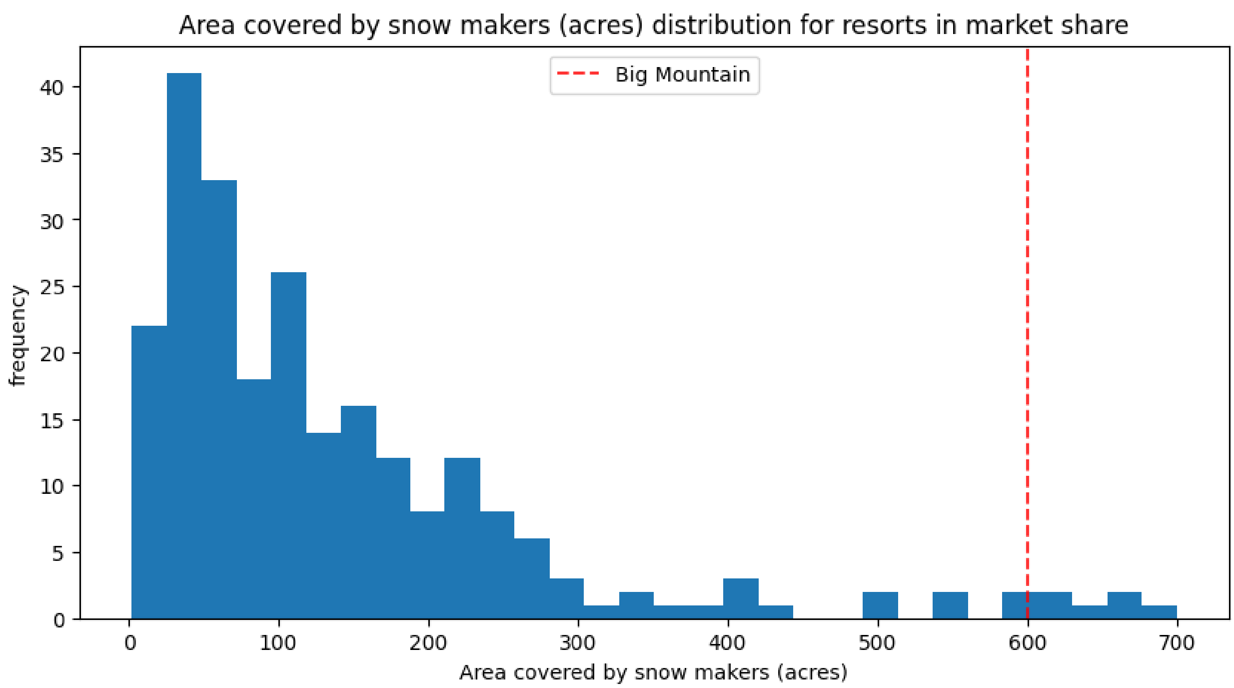
Winning model and scenario modeling

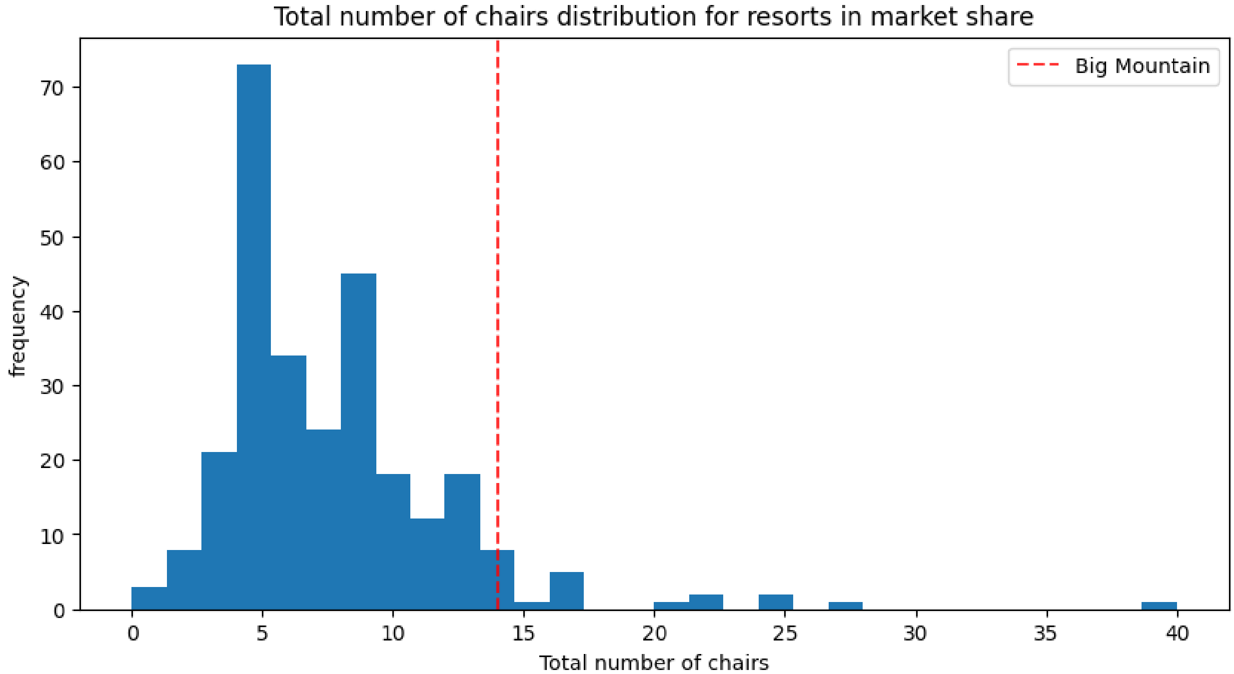
The random forest model is used to predict the ticket price for Big Mountain Resort.
Big Mountain Resort modeled price is $95.87, actual price is $81.00.

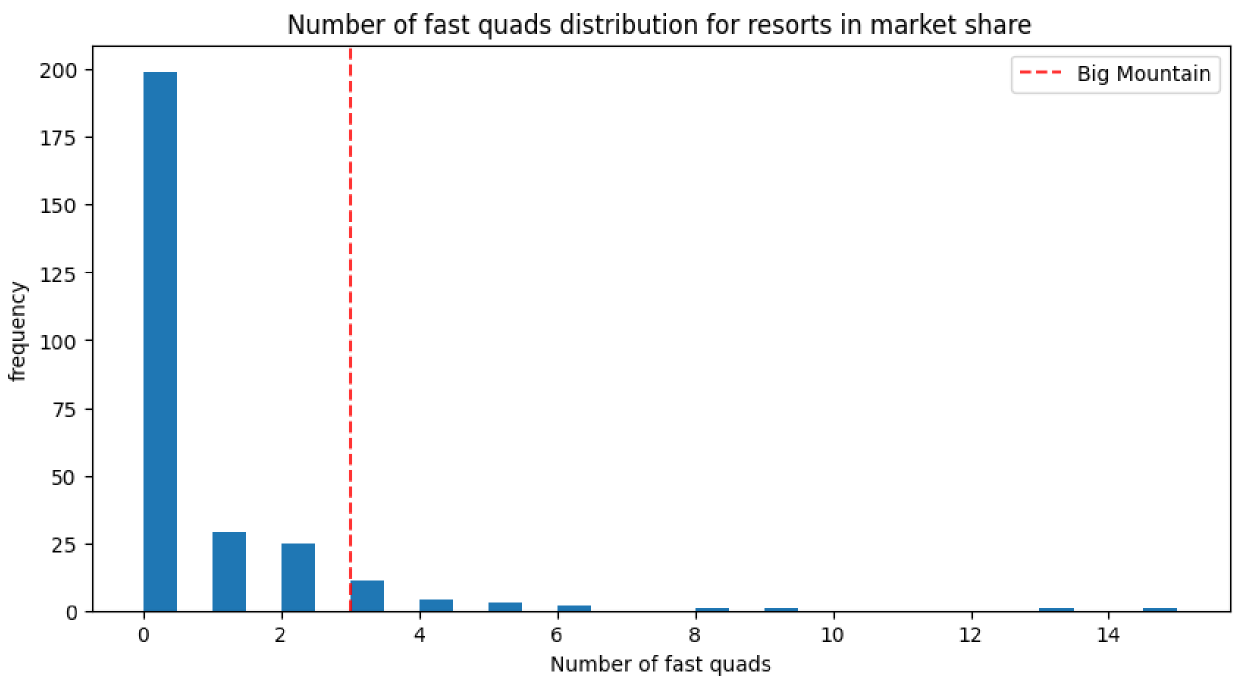
This model assumes that other resorts set their ticket prices based on the market demand. This assumption is not guaranteed. Further investigation is required.
From the linear regression model and random forest model, eight features that are most important to predicting the locket prices are vertical_drop, Snow Making_ac, total_chairs, fastQuads, Runs, LongestRun_mi, trams, SkiableTerrain_ac. Bar plots for these features and the ticket prices are used to compare those of the Big Mountain Resort with other skiing resorts.

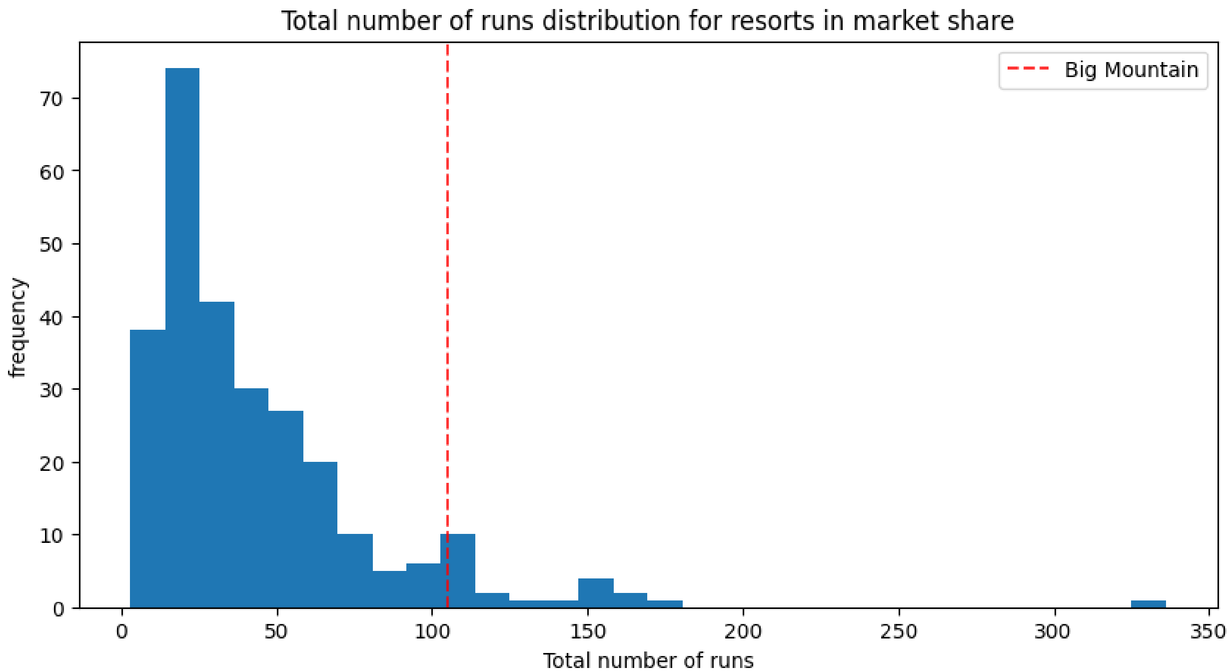


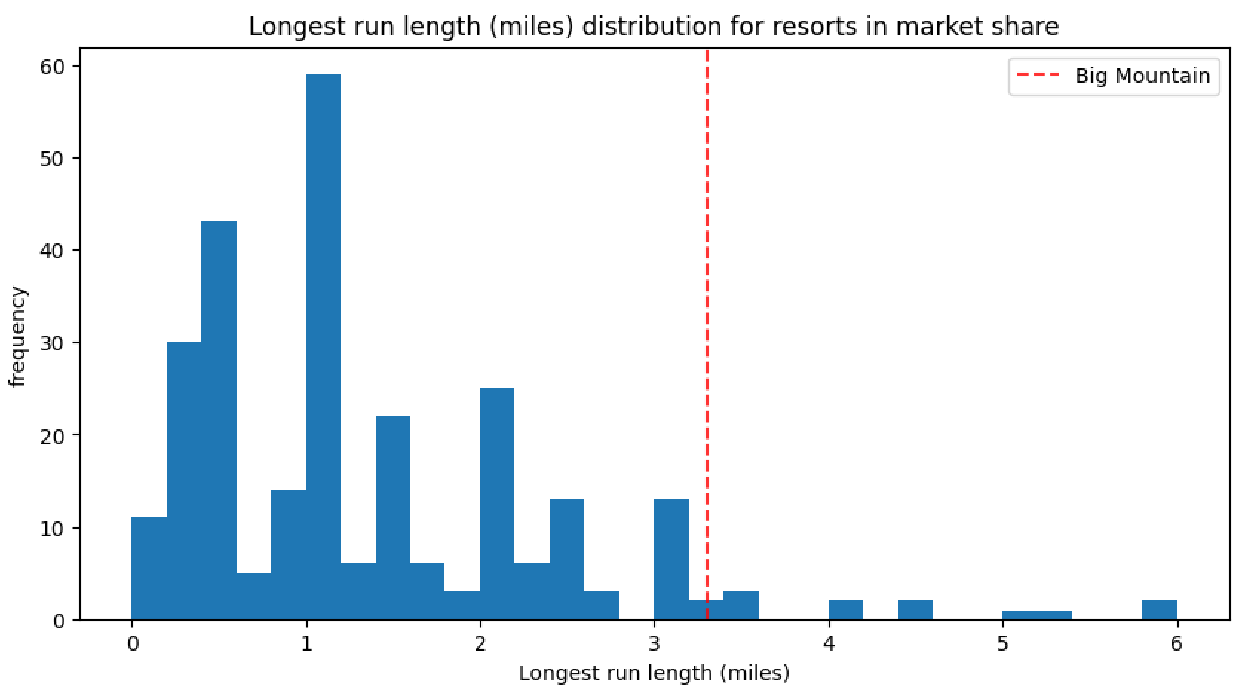

Vertical drop (feet) distribution for resorts in market share
Big Mountain
frequency
0
10
20
30
40
0
1000
2000
3000
4000
Vertical drop (feet)

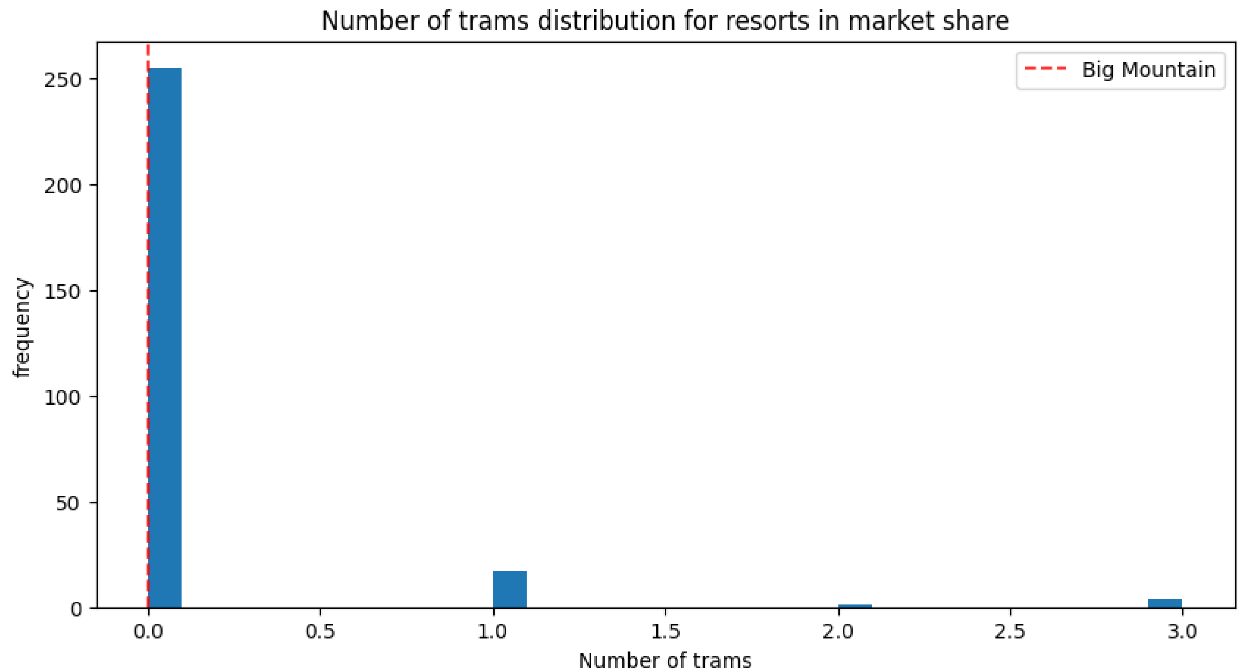

Area covered by snow makers (acres) distribution for resorts in market share
Big Mountain
frequency
0
5
10
15
20
25
30
35
40
0
100
200
300
400
500
600
700
Area covered by snow makers (acres)

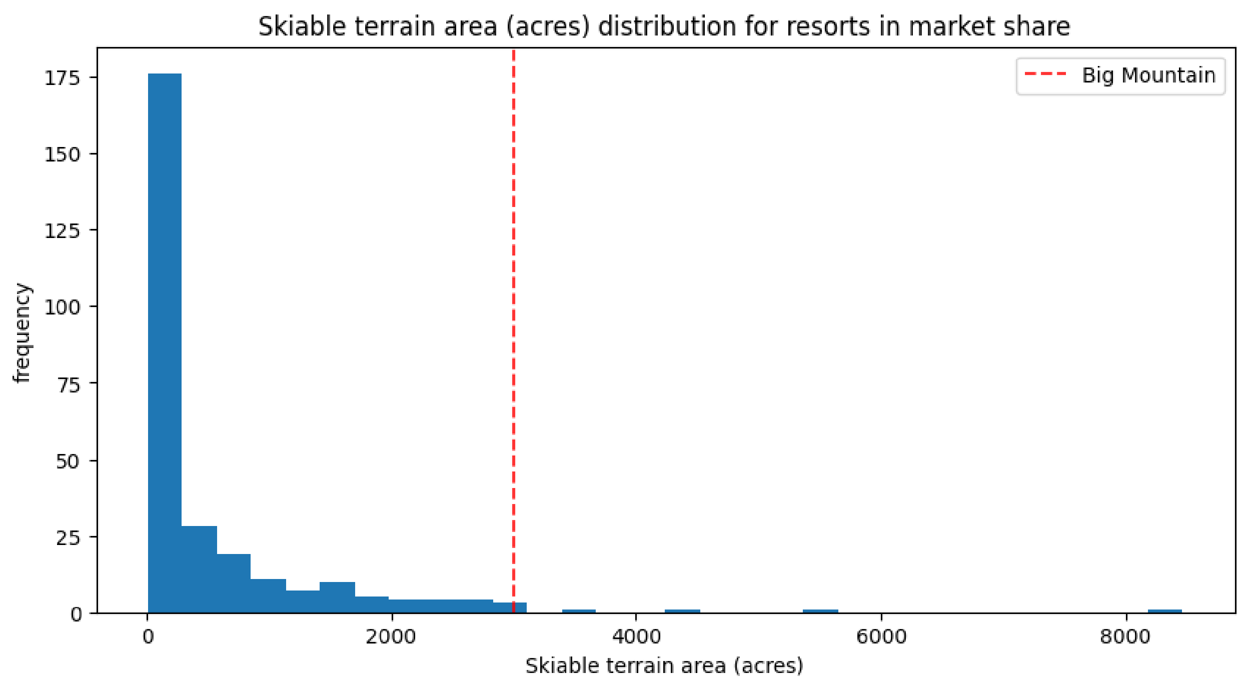

Total number of chairs distribution for resorts in market share
Big Mountain
frequency
Total number of chairs
0
5
10
15
20
25
30
35
40
0
10
20
30
40
50
60
70


Number of fast quads distribution for resorts in market share
Big Mountain
frequency
Number of fast quads
0
2
4
6
8
10
12
14
0
25
50
75
100
125
150
175
200


Total number of runs distribution for resorts in market share
Big Mountain
frequency
Total number of runs
0
10
20
30
40
50
60
70
0
50
100
150
200
250
300
350

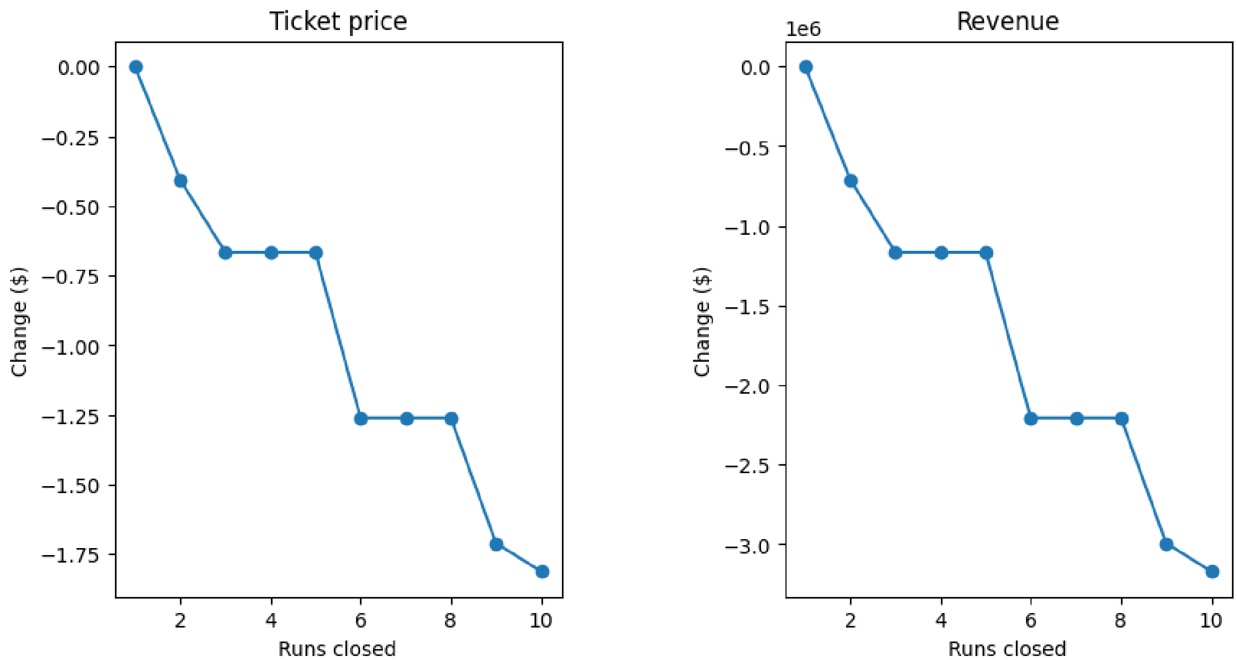

Longest run length (miles) distribution for resorts in market share
Big Mountain
frequency
Longest run length (miles)
0
10
20
30
40
50
60
0
1
2
3
4
5
6


Number of trams distribution for resorts in market share
Big Mountain
frequency
0
50
100
150
200
250
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Number of trams


Skiable terrain area (acres) distribution for resorts in market share
Big Mountain
frequency
0
25
50
75
100
125
150
175
0
2000
4000
6000
8000
Skiable terrain area (acres)

Four other pricing scenarios are analyzed to determine further adjustment in ticket prices.

1. Scenario 1: Closed up to 10 least used runs



   The line plots of Ticket price changes and Revenue changes with respect to runs closed from the model indicate that closing one least used run does not change the revenue and ticket price.

2. Scenario 2: Big Mountain is adding a run, increasing the vertical drop by 150 feet, and installing an additional chair lift. The model suggest an increase of $1.99 for ticket price, which increases the revenue to $3,474,638.00
3. Scenario 3: Big Mountain is adding a run, increasing the vertical drop by 150 feet, installing an additional chair lift and adding 2 acres of snow making. The model indicates that adding the 2 acres of snow making does not make any changes from scenario 1.
4. Scenario 4: Increase the longest run 0.2 mile and add 4 acres of snow making. The model suggest no additional ticket price and revenue increase

Conclusion

The current ticket price for adults on the weekend at Big Mountain is $$81.00. Our model suggests increasing the ticket price to $95.87 and still be competitive with respect to the state market. Further investigation suggests that additional chair lift operation cost could be offset by increasing the picket price by $1.99. Our suggesting model also indicates that closing one least used run will not have any effect on the total revenue, which suggests that one run should be closed to reduce the operating cost. Scenario 2 should be considered for future improvement. Adding one more run and increasing the vertical drop by 150 ft is supported by increasing the ticket price by $1.99.

Future scope of work

Our model is built based on the ticket prices of other resorts. Those prices, however, may not be set to reflect the true demand in the market. In addition, the resorts may not set their ticket prices based on the features in the model. Other operating costs that could be helpful to our models are the operation cost of the snow making machine, the labor cost required to operate the runs and skiable terrain. The current ticket price is lower than the model suggests could be due to the fact that the ticket price was set based on old data. The ticket price may not be updated with respect to the changes in the market. The business can use this model to update the ticket price every season. They can use this model to try other scenarios with other data set to improve the business.