Guided Capstone Project Report

Problem statement

Big Mountain Resort is a Montana based ski resort that offers eye-catching views of Glacier National Park and Flathead National Forest. Every year about 350,000 people ski or snowboard at Big Mountain. This resort is equipped with 11 lifts, 2 T-bars, and 1 magic carpet for novice skiers. In addition, the resort has recently installed an additional chair lift to help increase the distribution of visitors across the mountain. Due to the addition of new chair, operating cost has increased by $1,540,000 this season. Therefore, the management is interested to come up with better ticket pricing model for the ski resort. Apart from that, the management also want to identify the facilities for which the customers are likely to pay more for so that they can make future investments on those facilities.

Data wrangling

The data was available in CSV format in which rows have the name of the resorts, and columns have the name of the variables. In total, the dataset has information on 330 resorts across 38 regions (35 states) of the United States. Our primary target response feature is ticket price. However, there are two ticket prices- weekend and weekday ticket prices. Weekday and weekend ticket prices in Montana are equal. There were more missing values for weekday ticket price for Montana; therefore, weekday price was dropped and weekend price was considered as primary target response features. Target feature ‘fastEight’ was dropped from the dataset because majority of the data for this feature were missing. In addition, approximately, 14% of the rows have no price data and were removed from the dataset.

Exploratory Data Analysis

At the start of my analysis, I performed exploratory data analysis and visualization that allowed me to visualize the position of Big Mountain Resort in the overall ski resort market. State-wide summary statistics showed that Montana ranked fourth in terms of total skiable area although it was not among top ten in terms of number of resorts suggesting that Montana has fewer but larger resorts. In addition, Montano also ranked fourth in terms of resorts per capita.

This step also allowed me to get general idea about the variables that are associated with ticket price. AdultWeekend ticket price was positively correlated with several variables such as Snow

Making ac, resort_night_skiing_state_ratio, fastQuads, Runs, total_chair and Vertical_drop (**Figure 1**). We also checked the relationship between ticket price and total chairs to run ratio. When total chairs to run ratio was high, ticket price plummeted indicating that when the number of chairs is low relative to the number of runs, we can charge higher ticket price. But it should be considered that when the number of chairs is low, fewer number of visitors will be served affecting the total revenue. In this case, information on the number of visitors per year could have given more clear idea.

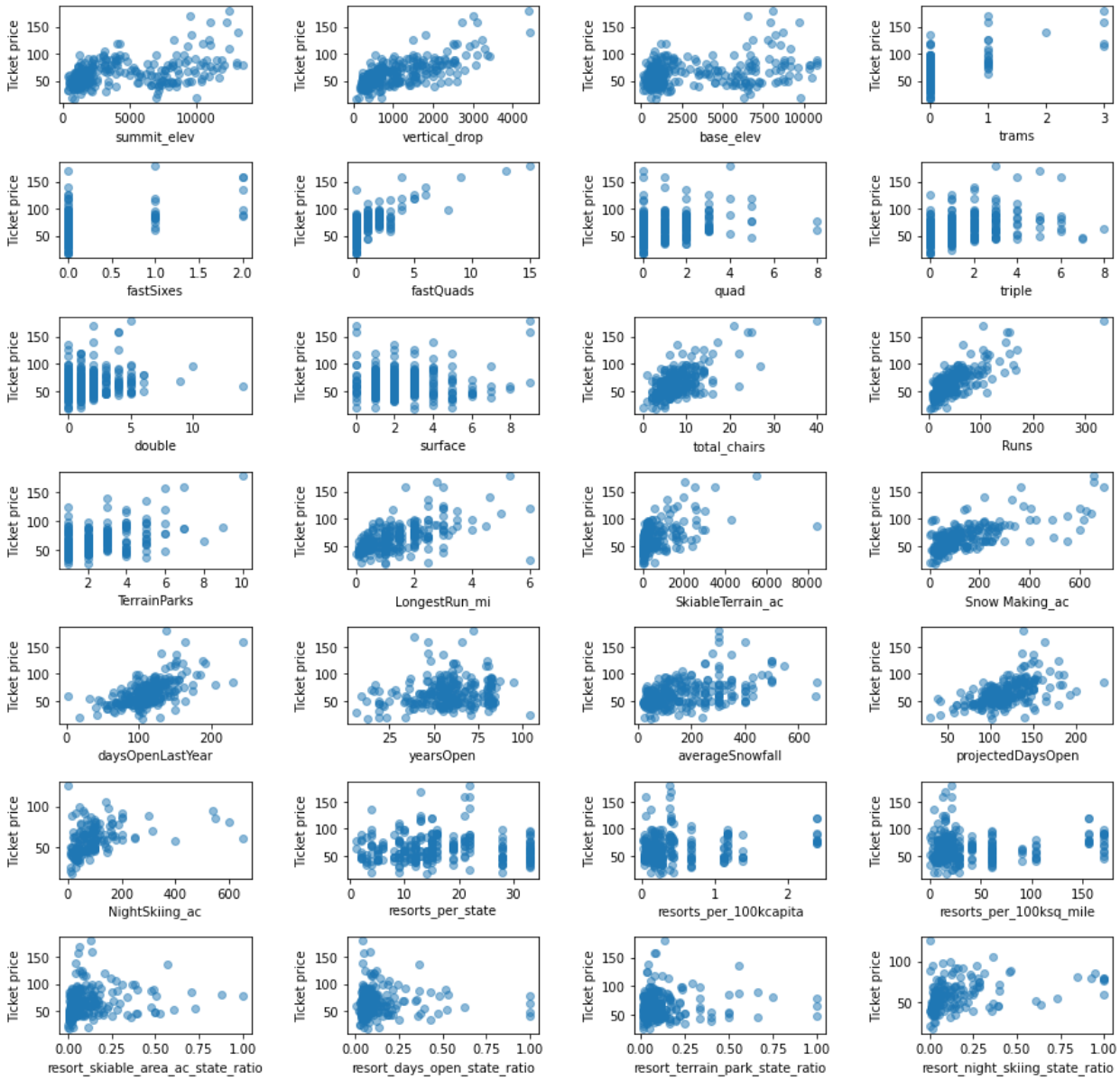


Figure 1. Scatterplots demonstrating relationship between features and ticket price

Modelling

I tested two different algorithms for this project: linear regression and random forests. A linear regression model was built and the model with cross-validation explained 63% of the variance in the ticket price. Likewise, a random forest model was also built and the model with cross-validation explained 69% in the ticket price. Cross-validation mean absolute error as well as the variability of random forest model was lower than that of linear regression model. Therefore, it appeared that model generated using random forests is better than the model generated using linear regression and I decided to use random forest model going forward. The top four features identified by both models included fastQuads, Runs, Snow Making_ac and vertical_drop (**Figure 2**). The learning curve function was used to assess if further data collection was necessary and it was found that no additional data is required.

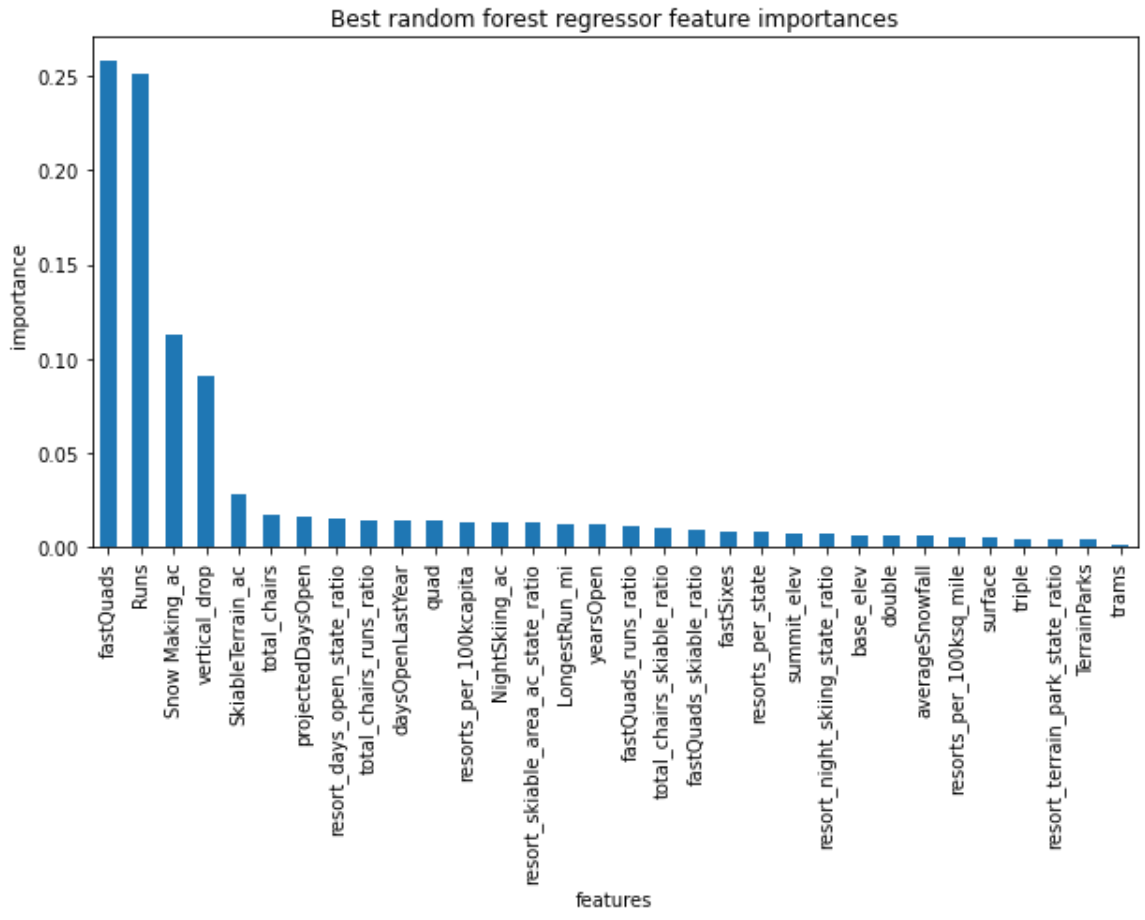
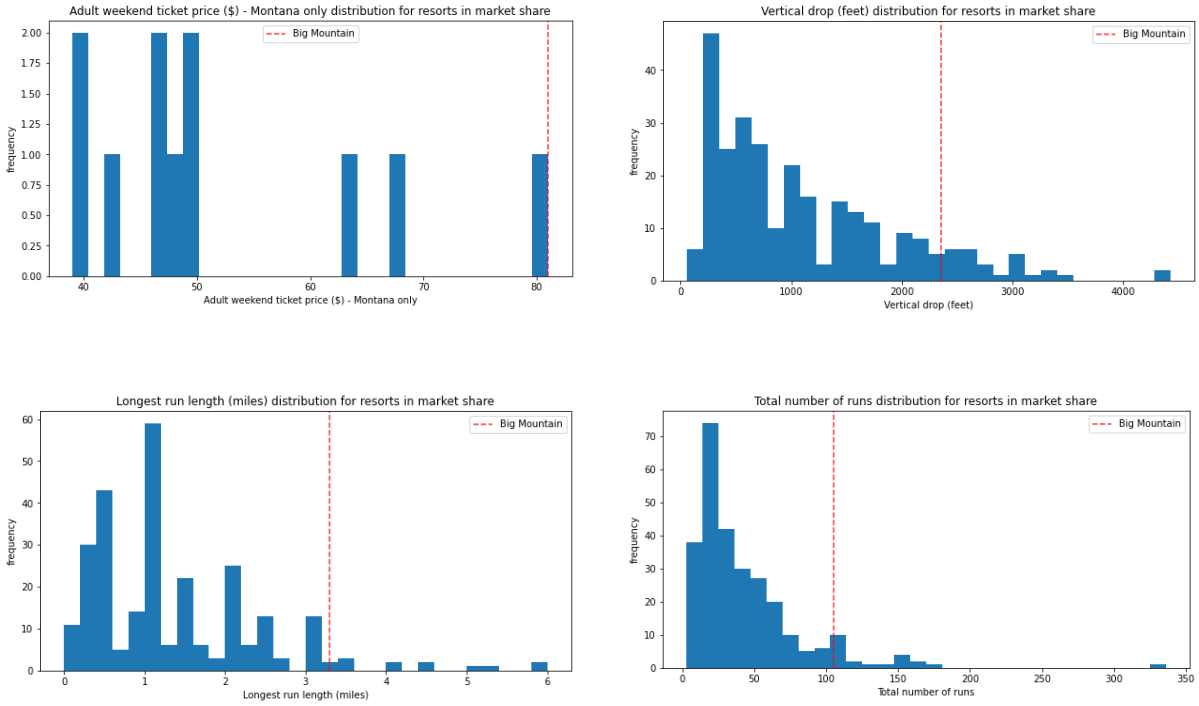


Figure 2. Significance of features identified using random forest model

Current ticket price of Big Mountain Resort is $81.00. Big Mountain Resort modelled price is $95.87 with mean absolute error of $10.39. Therefore, there is room for an increase in ticket price. Big Mountain Resort is amongst the resorts that rank top in the features such as vertical drop, snow making area, total number of chairs, fast quads, runs, longest runs, skiable terrain area which makes it one of the best resorts in the nation (**Figure 3**). But it should be noted that Big Mountain Resort already charges the highest ticket price in Montana. Since it is charging the highest ticket price in the state, I would recommend the management not to exceed ticket price above $85 (modelled price – mean absolute error). At the same time, I would also recommend the management to make other adjustments to reduce the operating cost. For example, the resort can close one run without any impact on ticket price while shutting down 2nd and 3rd run reduces support for ticket price. In case, third run is closed, 4th and 5th runs can be closed without any additional impact on ticket price and revenue (**Figure 4**).



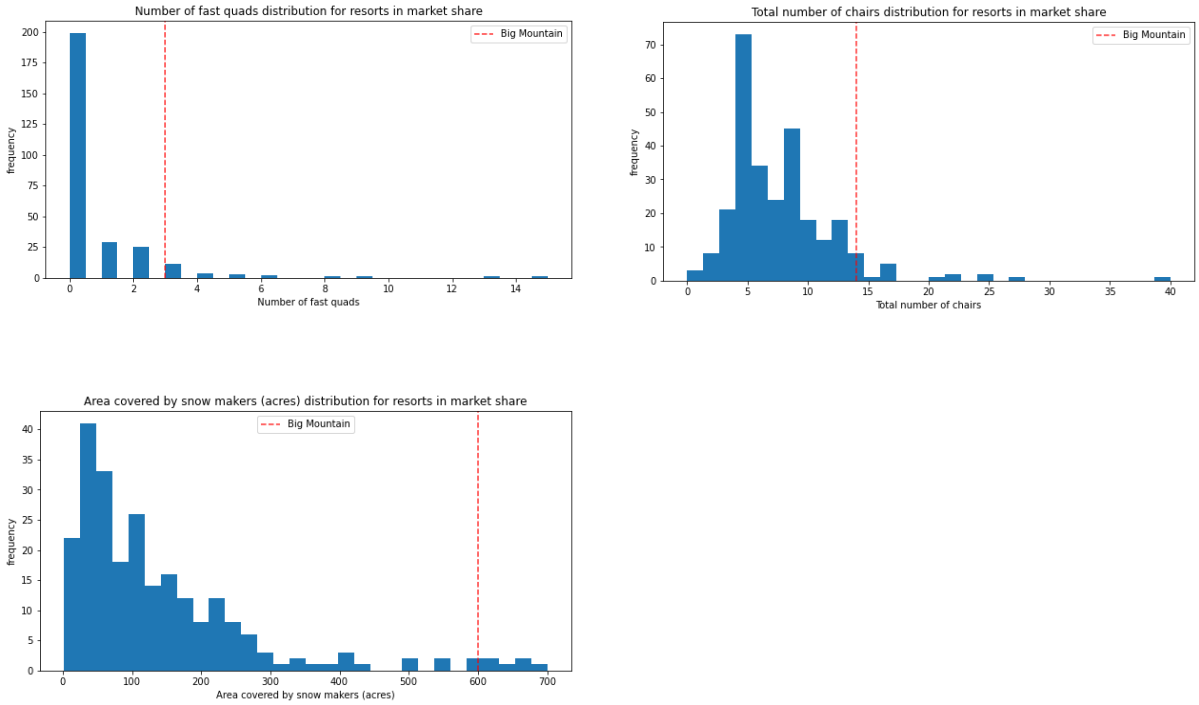


Figure 3. Position of Big Mountain resort

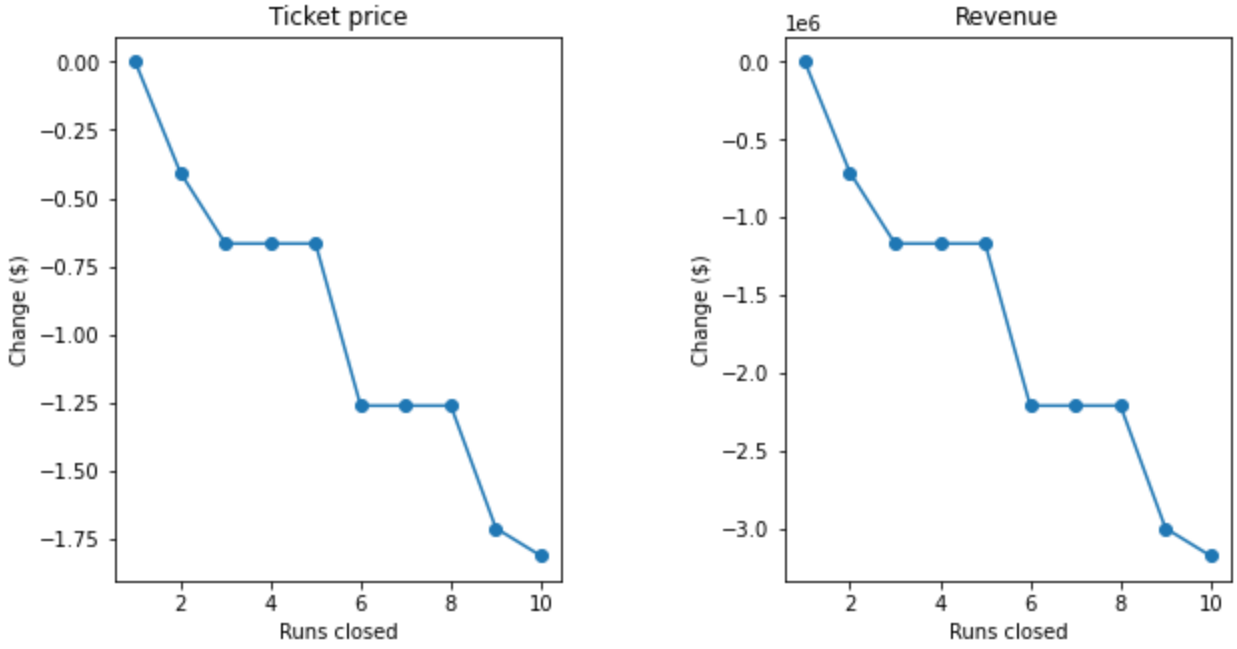


Figure 4. Impact of closing runs of ticket price and revenue