Guided Capstone Project Report
May 6th, 2024
Savannah Passaretti

Our client, Big Mountain Resort is a ski resort located in Montana with 105 trails that services 350,000 people per year. They installed a new chair lift that adds an additional $1540,000 to operating costs this season. BMR's current pricing strategy is to charge above the average of other resorts, rather than look at what facilities they are/aren't capitalizing on. To provide Big Mountain with a data-driven estimate on ticket pricing, the facilities need to be cost evaluated to determine which facilities they could emphasize or cut back on. Determining a ticket price that accurately reflects what the resort can offer will provide a more stable business strategy.

Given data on 330 different resorts, including Big Mountain, with 29 fields, we explored information we might want to use to investigate ticket pricing. We cleaned up the data that contained entries with null values or sometimes entire columns with insufficient data. After some exploring we noticed that Montana resort ticket prices are much different compared to the other states and weekday/weekend prices are 1:1. Ideally we would consult the client for insight on which markets we should specifically compare our resort data to, or if we should only look within Montana.

Moving forward, we explored certain features that looked troublesome and adjusted any outliers, etc. Using the correlation function in pandas, we investigated which features were highly correlated to ticket price. We plotted ticket price against other features, shown below:

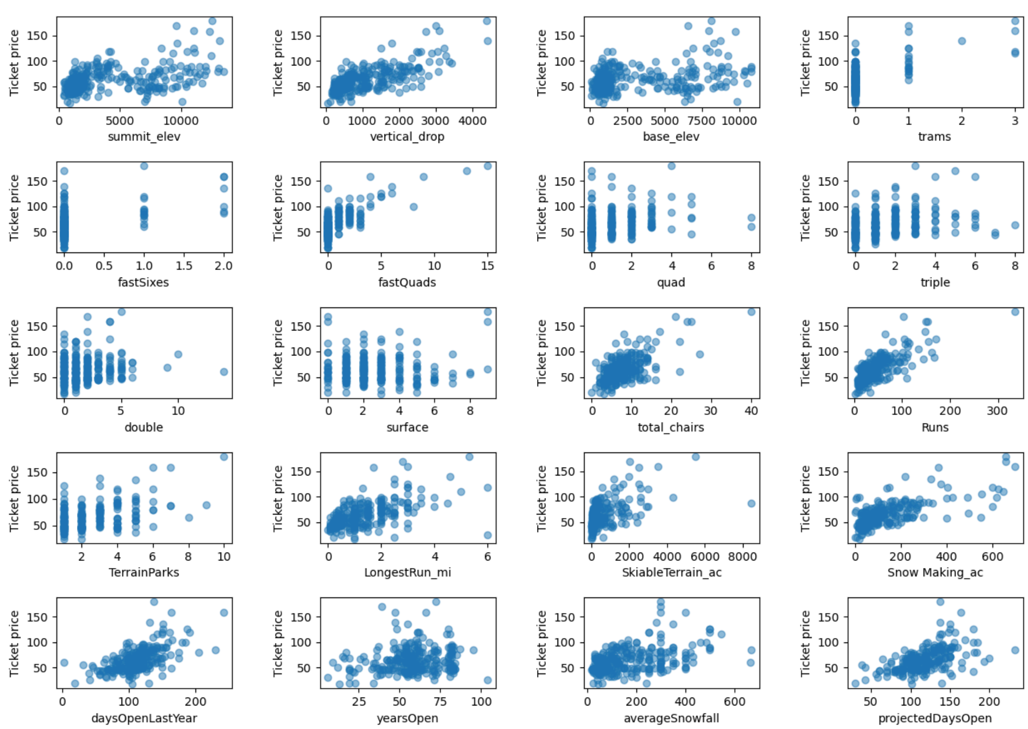

Figure 1: Ticket Price vs. Various Resort Features

The features that stand out are ‘fastQuads’, ‘Runs’, ‘Snow Making_ac’, ‘total_chairs and ‘vertical_drop’. After pointing out these key features we moved onto the pre-processing and training.

We split the data into two sets: Train and Test split 70% train and 30% test. We made predictions using an array of the mean of the training set and then used sklearn's Dummy Regressor to verify these would work the same. We continued to verify the training model by calculating several metrics like the coefficient of determination, mean absolute error and mean squared error. This is how we investigated the baseline of performance of the model.

The main models we explored were a random forest regression and a linear regression model. We found that the outputs of the random forest generator corroborate our top four features from the linear model (shown in Figure 2): ‘fastQuads’, ‘Runs’, ‘Snow Making_ac’, and ‘vertical_drop’.

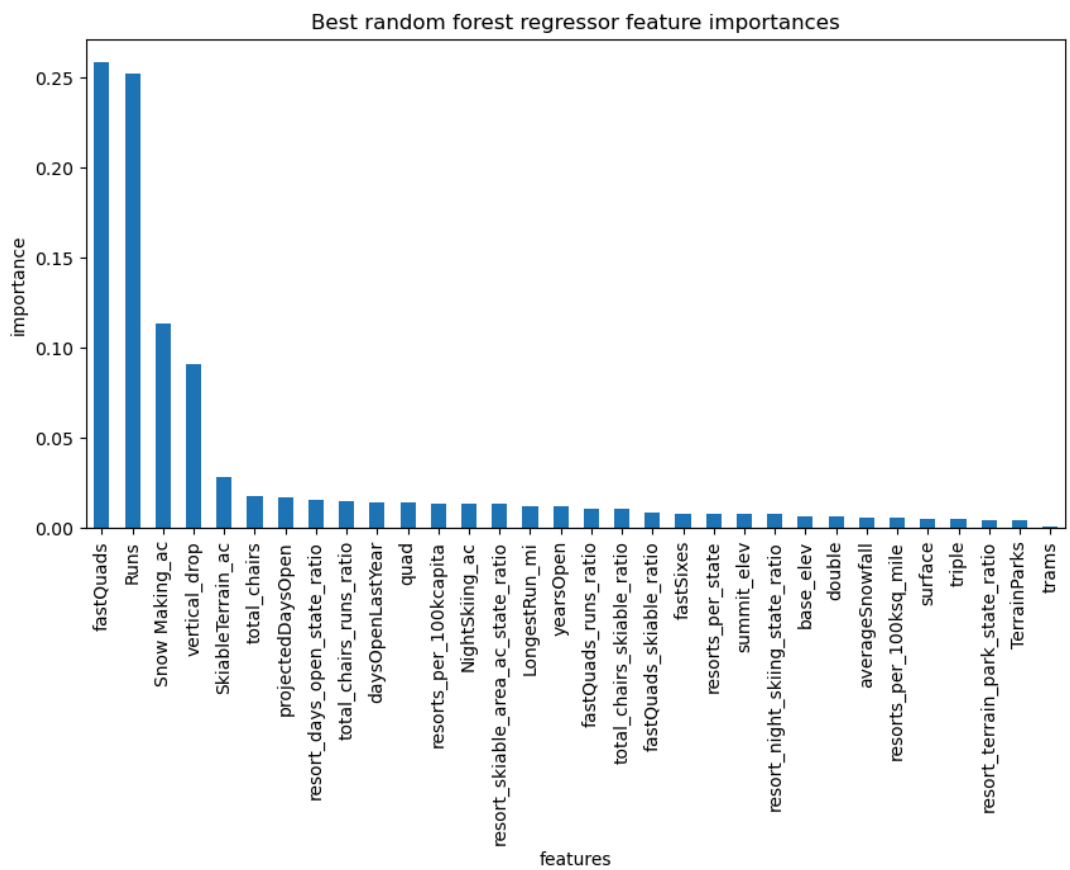

Figure 2: Random Forest Generator Feature Importance

Comparing the regression models the random forest generator showed less variability and after cross-validation showed a mean absolute error that is $1 lower than the linear regression model.

Big Mountain resort currently charges $81.00 for adult weekend tickets. Our model based on the ski resort marketplace suggests the resort should charge $95.87 per adult weekend ticket. This is assuming that the resorts our model is based off of are not overcharging or undercharging. We moved forward with resort characteristics to give more context to how different features support price increases or decreases. We first looked at how Big Mountain Resort compares to other resorts in terms of price, vertical drop, snow making, chair lifts, etc.

The potential areas Big Mountain may cut costs or increase revenue in are closing runs, adding vertical drop, adding snow making and/or increasing their longest run. Assuming there will be 350,000 visitors or ski on average 5 days a season, we ran some modeled scenarios to see how these characteristics might affect ticket price.

The most effective way to raise ticket prices would be to increase the vertical drop and add another chair lift to distribute skiers. We recommend moving forward with this scenario. Using similar modeling process, the resort can further test alternative scenarios and other variations to find how they would like to move forward.