Guided Capstone Report: Big Mountain Resort

Background

Big Mountain Resort, a ski resort in Montana with access to 105 trails, sees about 350,000 skiers and snowboarders a year. They have recently installed an additional chair lift to increase the distribution of visitors across the mountain. This additional chair has increased operating costs by $1.5M per season. Their current pricing strategy is to charge a premium above the market average of similar resorts, yet can set the ticket price freely.

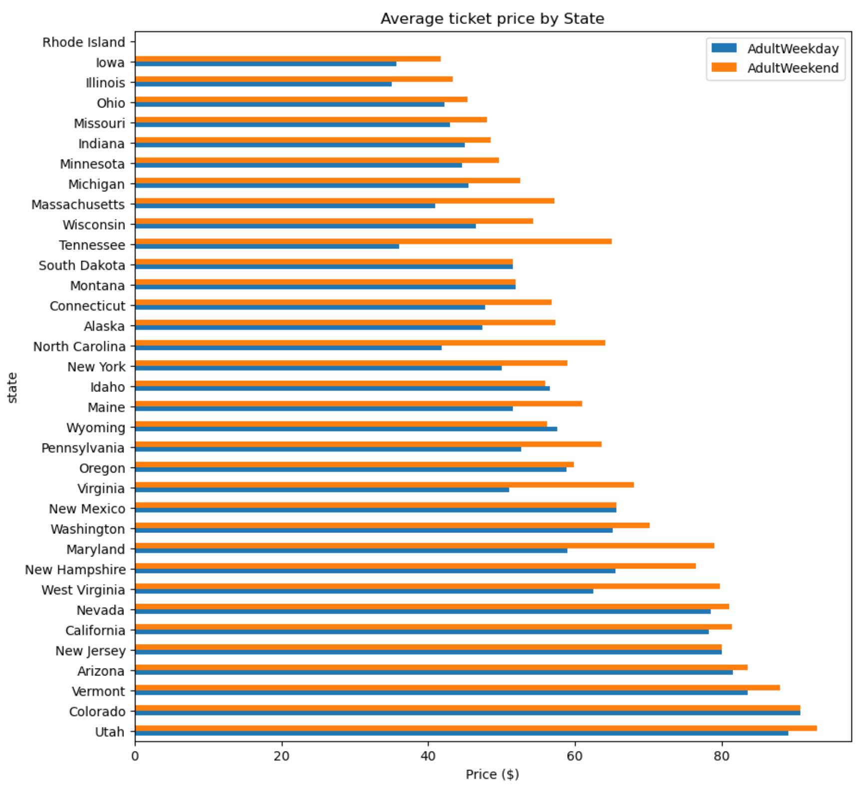
Problem Identification

What steps can Big Mountain Resort take this ski season to offset the increase in operating costs. There is concern that BMR is not capitalizing on all of its facilities. The leadership would like guidance and insights to improving revenue through their ticket pricing model without raising operating costs.

Data Wrangling

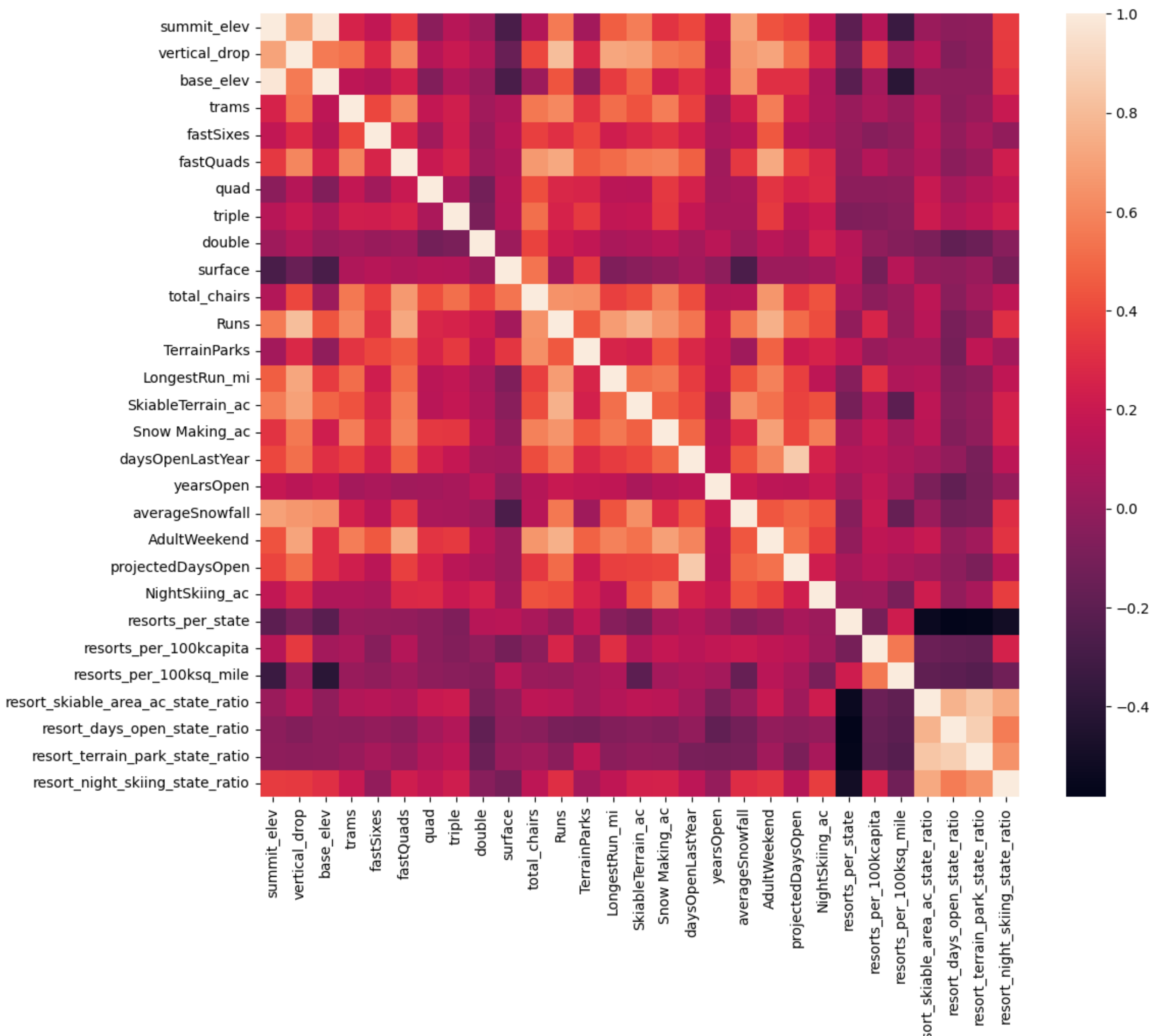
The project aims to build a predictive model for ticket price based on a number of facilities or properties boasted by resorts. The goal of this data wrangling step is to identify the required target value and any potentially useful features. It is also important to locate any fundamental issues with the data and correct it.

The data we used has info about 330 resorts in the US. There are twenty-seven features, two of which are weekend and weekday prices, and the rest are information about the facilities at the resorts. The weekend vs weekday prices were plotted as below (Fig.1) The target value we identified to model resort ticket price on is ‘AdultWeekend’, while 'AdultWeekday' was omitted as it contained more missing values than weekend prices. The final dataset was truncated to 227 rows across 25 columns after the records with missing values were dropped.



Exploratory Data Analysis

As a result of the wide variability in the ticket price within some states and between the states themselves, the data was augmented with state population and area information from another dataset to draw some state-wide summary statistics. This allowed us to formulate 2 new variables to measure resort density. A heat map (Fig. 2) and series of scatter plots were constructed to visualize correlations between different numerical features and price data. Here we were able to identify several potential factors that led to higher ticket prices to be further investigated in the next step. These include: vertical_drop, fastQuads, total_chairs, Runs, Snow Making_ac, and NightSkiing_ac.

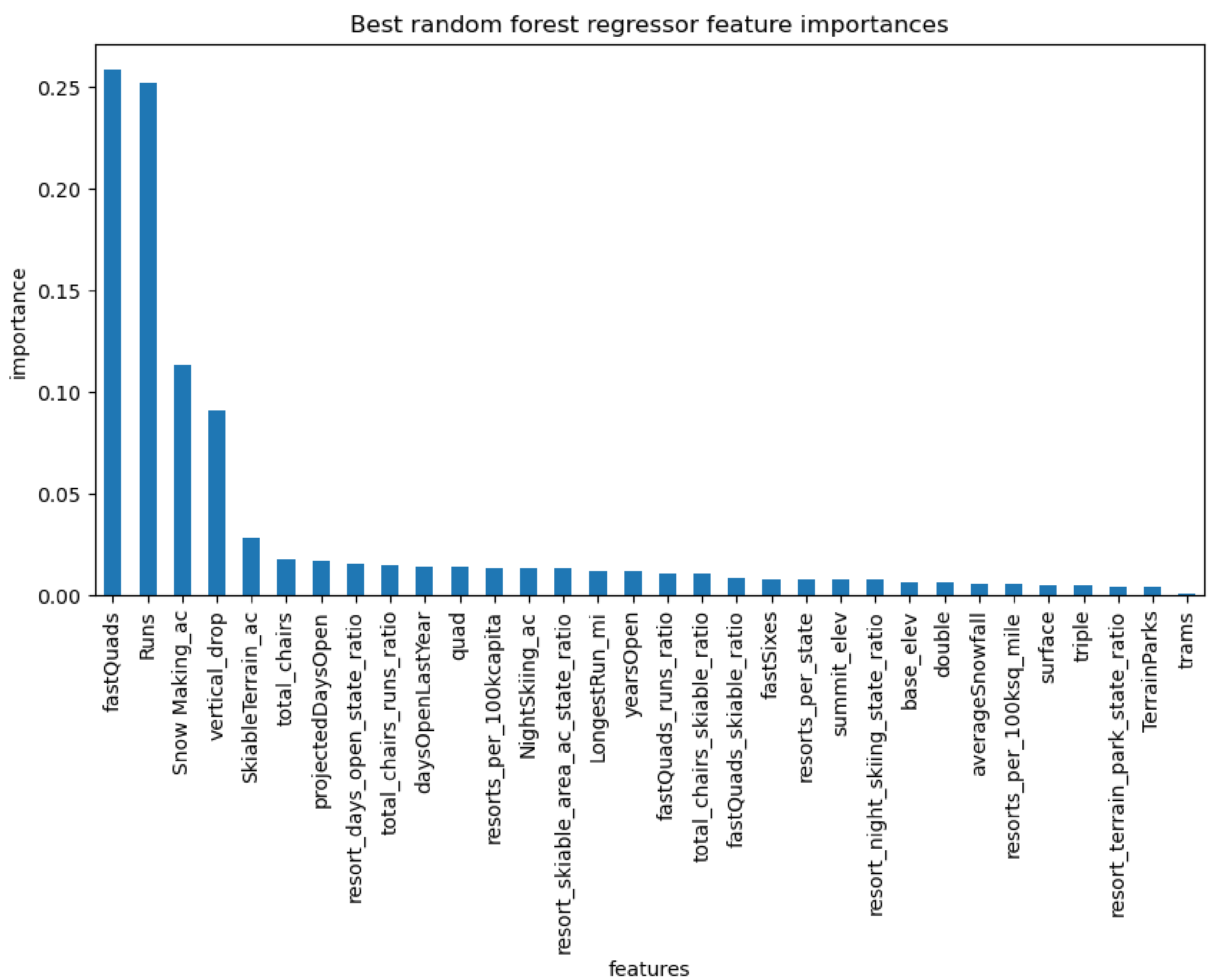


(Fig. 2)

Model Preprocessing with feature engineering
The first step of developing our model was examining performance based on the average price. This proved to be useful in establishing a baseline for comparison. Utilizing the MAE, we found the predicted ticket price by using this method, on average, might be off by around $19.

Moving on, we decided to build a linear model. The first step was to impute the missing values with median and mean values. The results indicated that in both cases, the ticket prices would be off by about /$9. While this is better than estimating using the average, we suspect the model may be overfitting. By using a technique called cross-validation, we are able to partition the training set into folds, and calculate performance on the fold not used in the training. We then visualized these most useful features (vertical_drop, Snow Making_ac, total_chairs, fastQuads, Runs, LongestRun_mi, trams, SkiableTerrain_ac), and this was consistent with our assumptions made during EDA.

In addition to the linear model, a Random Forest Regressor was built, also with missing values imputed using the mean and median values. The performance of this model was also assessed using cross-validation. The dominant top 4 features revealed were in common with the Linear model (Fig. 3).



Guided Capstone Report: Big Mountain Resort

Algorithms used to build the model with evaluation metric
We compared the two models' performance. The random forest model has a lower cross-validation mean absolute error by almost $1. It also exhibits less variability. Verifying performance on the test set produces performances consistent with the cross-validation results. As such, we will use this model. Data quantity assessment was performed using the learning_curve function, and concluded that there is enough data.

Winning model and scenario modeling
BMR has been reviewing potential scenarios for either cutting costs or increasing revenue. The following options have been shortlisted:

1. Permanently closing down up to 10 of the least used runs. This doesn't impact any other resort statistics.
2. Increase the vertical drop by adding a run to a point 150 feet lower down but requiring the installation of an additional chair lift to bring skiers back up, without additional snow making coverage
3. Same as number 2, but adding 2 acres of snow making cover
4. Increase the longest run by 0.2 mile to boast 3.5 miles length, requiring an additional snow making coverage of 4 acres

We don't have information on the runs' usage, but the model says closing one run makes no difference while closing 2-3 reduces support for ticket price and revenue. It also shows that if closing down 3 runs, closing down an additional 4th and 5th will have no further loss in ticket price. However, increasing the closures to 6 or more leads to a large drop.

In the second and third scenarios, the model increases support for ticket price by $1.99 which could be expected to amount to $3,474,638 in revenue over the season.

The final scenario showed no difference in price according to the model.That is when they decided to appeal the process.

Pricing recommendation
Resort ticket price modeling relies on the assumption that other resorts are setting the price based on what people value at certain facilities. Currently, the price of an `AdultWeekend` ticket at Big Mountain is $81.0. We visualized where it sits overall among all the other resorts in the market share by plotting the price distribution of the resorts in the data (Fig. 4). According to the model, BMR could charge $95.87 and be supported in the marketplace; even an expected mean absolute error of $10.39 indicates room for an increase. An increase in ticket price would help pay for the operating costs of the new chair lift.