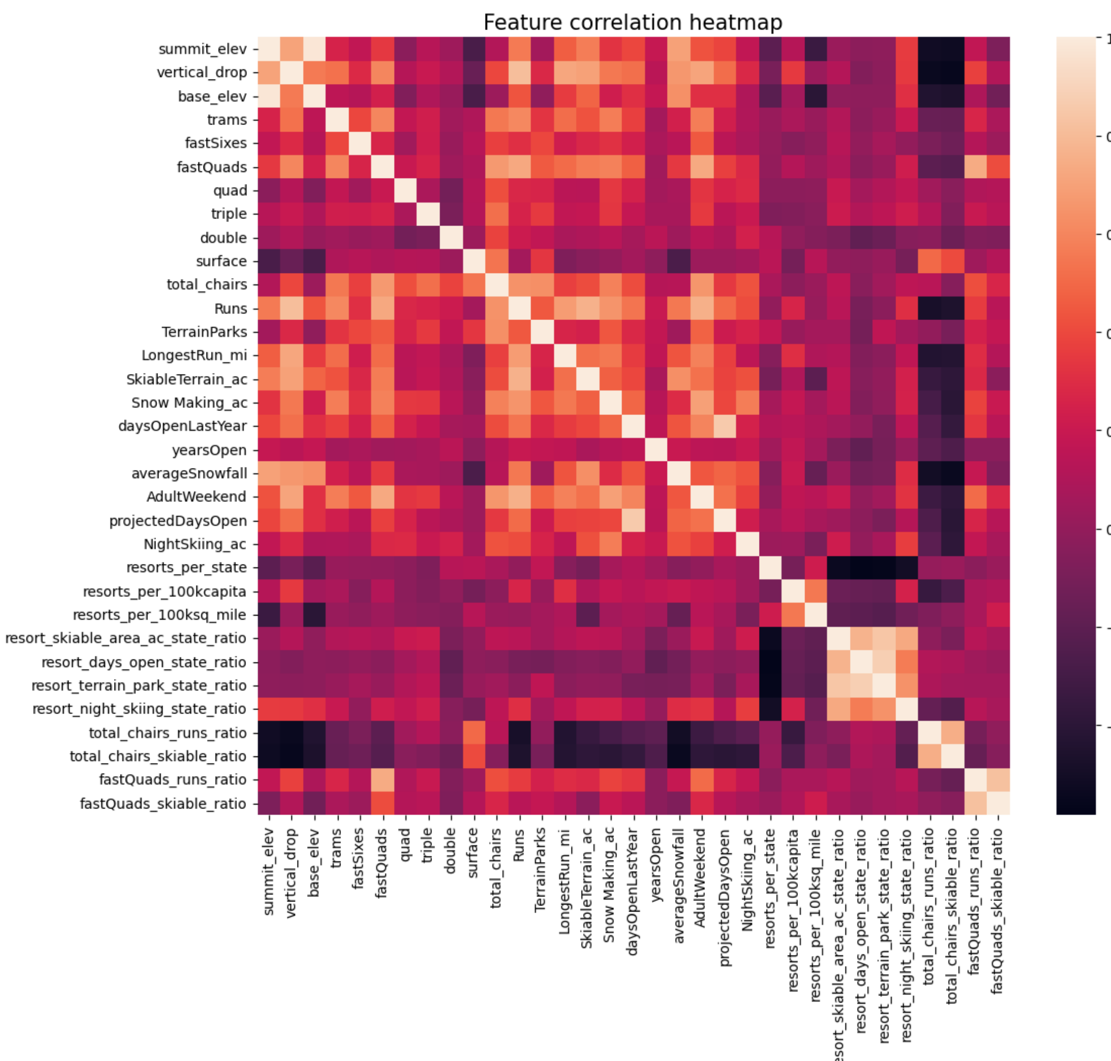
Big Mountain Resort (BRM) is reassessing its ticket pricing strategy after investing in a new chairlift that increases operating costs for the upcoming season by about $1.5 million. BMR's current ticket price is $81.00 and is based on the average ticket price within its market segment plus a premium to account for the resort's added facilities. The goal of my analysis is to propose an optimized pricing strategy that will offset the additional operating costs using a model that identifies key features and facilities that are the most valuable to visitors.

I utilized the dataset provided by the Data Manager of resorts within BMR's market segment. The dataset encompasses categorical features such as state and region, as well as numeric features including elevation, number or different types of chairlifts, number of runs and terrain parks, park areas, days open, night skiing, and ticket prices most importantly. After an initial evaluation, I was able to clean the data. There were also a number of resorts were missing ticket price information, which had to be removed from the dataset because ticket price is our target feature.

After combining the resort data with pertinent state population information and ratios, there were not apparent grouping nor outliers of states when determining ticket prices. Therefore, I concluded that states can be treated equally in this analysis The heatmap below shows the correlation of numeric features of the data. Particularly focusing on ticket prices, *AdultWeekend*, we can see that the lighter colored squares have a higher positive correlation such as vertical drop, fast quads, runs, and snow making acreage. We can predict that these variables will be important when modeling BMR's ticket prices.



Moving on to modeling the data, I considered three models: mean, linear regression, and random forest regression. Each model was trained on 70% of the data, then tested and evaluated on the remaining 30%. As the name implies, the mean model predicts the ticket price based on the average of each feature. Next, the linear regression model was build using a pipeline that imputes missing values with the median, scales the data, selects the best features, and trains the model.

The linear model was initially overfitted to the training data when using all available features. A grid search cross-validation revealed that using 8 features was optimal for this model, which are vertical drop, snow making acreage, total chairlifts, fast quads, runs, longest run (miles), trams, and skiable terrain. Finally, the random forest regression model has a similar pipeline as outlined above. I compared results using various assumptions, and concluded the best random forest regression model would have 69 trees, impute with the median value, and have no scaling. Comparing the most important features of the random forest regressor showed four features with substantially higher importance than the rest: runs, fast quads, snow making acreage, and vertical drop. These four features are the same as those highlighted by the heatmap above.

With these three models finalized, I tested each on the remaining 30% of data and compared the mean absolute error (MEA) to conclude which model is best, results below.

Model	MEA
Mean	19.13
Linear Regression	11.79
Random Forest Regression	9.54

We can see that the random forest regression model has the lowest MEA, and therefore is the model I selected to analyze ticket price. At this point I can refit the random forest regression with all of the data, excluding BMR, then calculate the expected ticket price for BMR. The modeled price is $95.87, compared to their actual ticket price of $81.00.

I recommend BMR increase their ticket price to the modeled $95.87. This would well exceed the recent addition operating cost of the new chairlift, and the price is a better reflection of the value of BMR’s facilities. Another BMR can consider is increasing the resort’s vertical drop by adding a new run and chairlift. With this addition, the model supports increasing ticket prices by only $1.99 and is associated with $3.47 million of additional revenue during the season.

In conclusion, BMR’s current ticket price undervalues the resort, and there is substantial room to increase prices and revenue. This model is now available to BMR so that they can test their own assumptions and see the impacts of various scenarios in real time. In the future, having additional data on visitor counts for all resorts would be useful to gain a better understanding of the relationship between price, number of visitors, and to what extent the facilities are being utilized at capacity. This model could be augmented to account for this visitor information and run sensitivity analyses.