Problem Statement: What options are there, in terms of a pricing strategy, that will allow Big Mountain Resort to increase revenue by at least $1,540,000 this upcoming winter season?

Big Mountain Resort is a ski resort located in Montana. They service about 350,000 skiers/snowboards who spend 5 days, on average, at the resort, at $81 a day. Big Mountain recently installed a chair lift to help increase the distribution of visitors across the mountain. The additional chair is projected to increase operating costs by $1,540,000 this upcoming season. Big Mountain’s pricing strategy has historically been to charge a premium above the average price of resorts within their market segment. The company is suspicious that they may not be capitalizing on their facilities as much as one could be and are seeking other strategies to account for the importance of some facilities they have.

Data Wrangling/ Exploratory Data Analysis/Feature Engineering:

The data acquired was packaged in a CSV file from Database Manager provided by Big Mountain Resort about resorts within the US in Big Mountain’s market. Features included pricing data and features of each resort like region, state, number of quads, and skiing area. Preliminary analysis of the data involved sorting through missing/incorrect values, determining if duplicate results were present, and categorizing resorts by state. Since pricing data was subdivided into weekday and weekend data, and a larger percentage of data was missing by weekday, weekend data was prioritized for data analysis in future steps and used to extrapolate weekday prices. To validate this assumption, data related to weekday and weekend data was grouped by state to evaluate their distribution. We were able to conclude that prices, in general, are only differentiated by day of the week if prices of the tickets are less than $100.

Exploratory Data Analysis/Feature engineering:

Data was then aggregated based on state to determine if market differences were present by state. This includes features such as resorts/state, total skiable area/state, and total days open per state. After scaling, variance analysis concluded that 2 features accounted for 77% of the variance in variance pricing, which were resorts per 100k capita and resorts per 100k sq miles.

Chart, scatter chart

Description automatically generatedThis data was plotted with the combined two features vs average weekend prices to determine if there was a correlation present. An obvious correlation was unable to be determined when plotting these results. Next, features were created for each resort on a per state basis to determine ratios of how much market share resorts had within their state (ex. Resort terrain park per state). A heat map of the correlation between features, with a particular interest to adult weekend price, was created to visual results.

Fast Quads, Runs, Snow making, and resort night skiing state ratio had the highest correlation with weekend prices. *These plot is show to the left.*

Model Processing:

A training (70%) and test set (30%) was created from the resorts available excluding ‘Big Mountain’. The prediction was for weekend price. Features in the training and test data with missing values were replaced with the mean and median and stored separately. The first pipeline used an input strategy of filling in missing features with the median result, using a standard scaler, and selecting k=10 best features with a scoring function of f\_regression (the cross correlation between each regressor and the target is calculated to get an r regression and converted to a f-score and finally a p-value). Finally, a linear regression is created with the selected features. The mean absolute error was 9.501 for the for the training set and 11.202 for the test set. This was worse without selecting k features. K was increased to 15 with slightly better results, but MAE for both the training and test set did not perform as well as the original linear regression model with median missing features.

Algorithm used to build models and evaluation metric:

Chart, histogram

Description automatically generated5-fold cross-validation was done different pipeline models. The mean cross validation score was 0.633 with a standard deviation of 0.095. k select models were created (max of all features collected) followed by 5-fold cross validation to evaluate performance. The best performance was with k=8 features selected: vertical\_drop, snow\_making\_ac, total\_chairs, fastQuads, Runs, LongestRun\_mi, trams, and SkiableTerrain\_ac. The MAE for this model on the test data was 11.783

A random forest regressor model was created as well. 5-fold cross validation was done to assess performance. The best model had n estimators = 69, median missing data strategy, and no standard scaler used. The MAE for this model was 9.537 on the test data. The random forest model was saved. *This models features is located to the left.*

Finally, a data quality assessment was done. A learning curve was created by varying the train-set size. A leveling off occurred around 40-50 data points for our model, which is significantly less than the amount collected.

Pricing Recommendation, Conclusion and Future Work:

The potential scenarios to evaluate including, (1)permanently closing down up to 10 of the least used runs, (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 with the addition of 2 acres of snow making cover, and (4) increasing the longest run by 0.2 mile to boast 3.5 miles length, requiring an additional snow making coverage of 4 acres. **Scenario 2 appears to be most the favorable options in terms of revenue, supporting a price hike of $1.99 and resulting in an $3.47 million.**

Findings support a price hike up to around $95 without any change implemented. Assuming visitor numbers stays the same at 350,000 per year, and 5 tickets purchased per visitor, this will result in a revenue increase of around $24.5 million. However, if our client is worried about a significant price hike from $81 in Montana, where they are already the most expensive option. There were certain limitations to the work done. Follow up with executives and historical pricing data would be prudent to access. There are many aspects to a pricing strategy. For example, do we want to charge per day, or charge per feature? Do we have different bundles that we could offer? What about other aspects of the business that create revenue and how will increasing admission affect other revenue generated? Discussing marketing strategies with the marketing team seems most prudent. Follow up with customers and how they may respond to price changes through a customer survey may provide insight, as demand for Big Mountain may be affected.