Big Mountain Resort Project Report

In this report, we used the data science method to help identify our problem, gather and understand the data as well as model it to help visually understand how to come to a solution and put it into a summarized report.

Step 1: Problem Identification

Big mountain resort plans to charge higher premiums on their ticket and predict they are not capitalizing on the facilities they provide. What ways can Big Mtn Resort maximize their revenue by optimizing ticket prices that give value to the customers as well as compare facilities with other resorts in their market using data-driven strategies?

Step 2: Data Wrangling

In this step we took the data given and did steps to give us cleaner data that can provide more data quality and usability. In the data we found some missing values and duplicates that had to be cleaned up. In addition there were also outliers that had to be examined to see if they needed to be corrected or removed. There were also data that had to be converted into a more consistent and clear format.

Step 3: Exploratory Data Analysis

In this process we summarize key stats then convert the data into visuals where we can help identify any trends of relationships in the data.

We found that the numerical/categorical features that were in the data was the ski_data set that contained 277 entries and 25 columns which consisted for features such as Region and states. There appears to be no relationship between states and ticket prices. The features that are important for the ticket pricing modeling would be Runs, total_chairs, Snow Making _ac, vertical_drop, and fastQuads.vertical_drop, and fastQuads.

Step 4: Pre-processing and Training Data Development

In this step we took data and applied different machine learning models to help identify what model is best to implement for the resort.

One of the initial steps of process was to partition data using the Train/Test split which helps to show the expected model future performance. We also applies cross-validation technique to help determine which features generate the best performance. To assist with making things more simplified and to see what model is best used for modeling we incorporated the linear and random forest regression model performance.
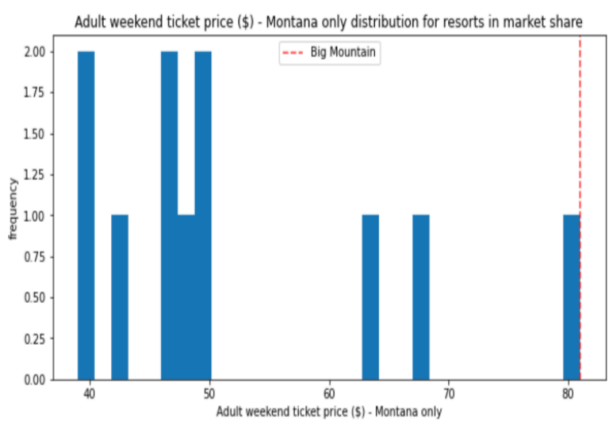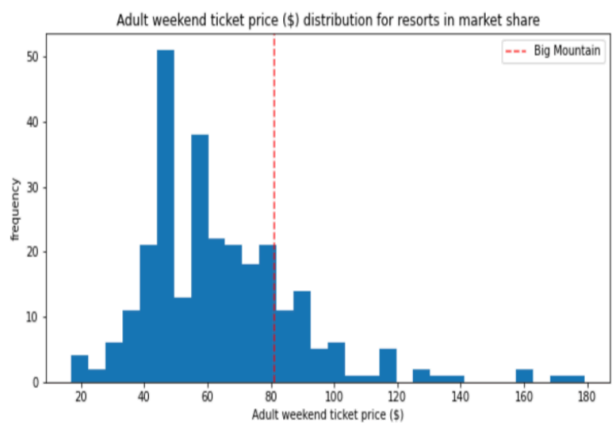In conclusion, we use the learning-curve function to see if business needs to undertake further date collection.

Best random forest regressor feature importances

Cross-validation score as training set size increases
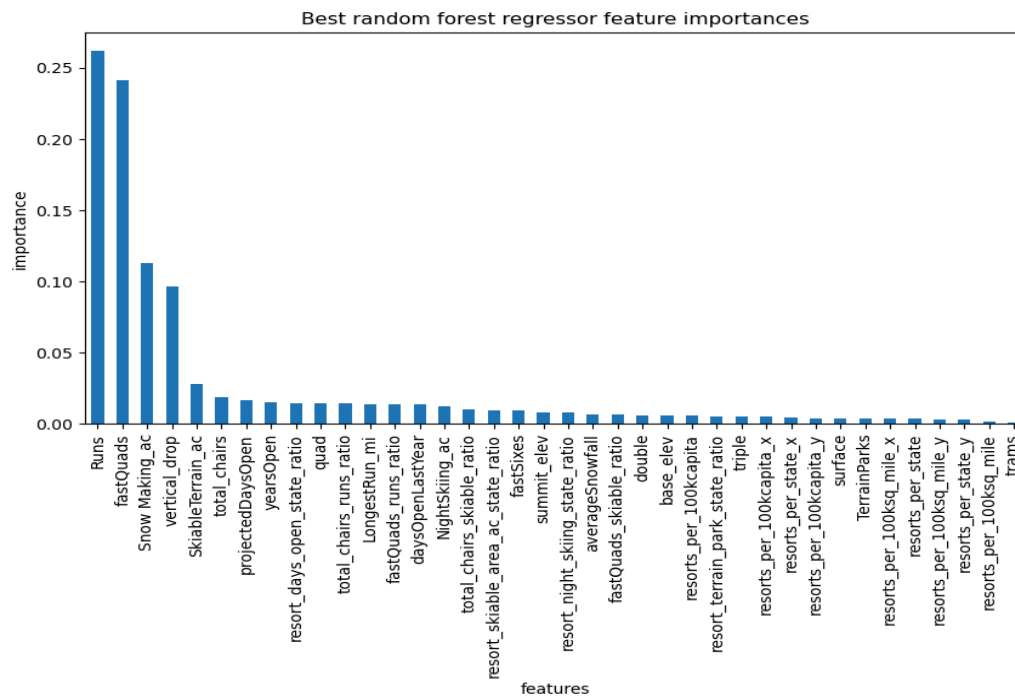
Step 5: Modeling

This step involved choosing the best model and taking the data and further mapping and comparing the ticket prices of Big Mtn with other resorts as well as in Montana only.

We compared the different features of Big Mtn Resort with other resorts to see how Big mtn is performing overall.
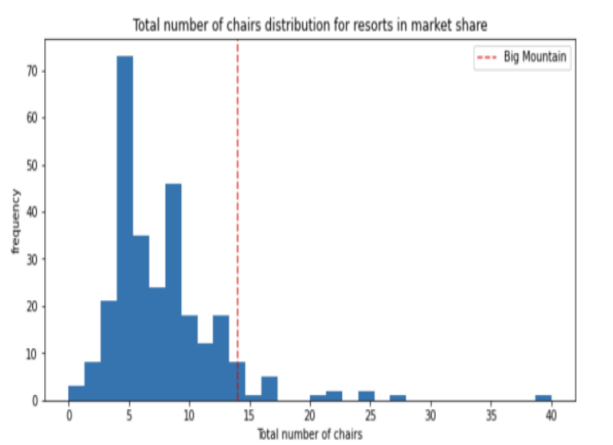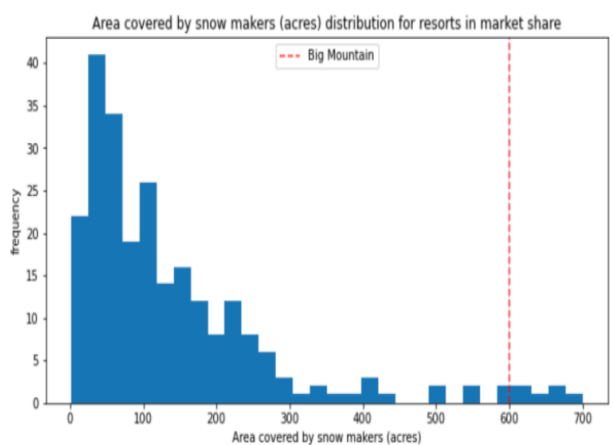

Adult weekend ticket price ($) distribution for resorts in market share

Adult weekend ticket price ($) - Montana only distribution for resorts in market share

Snow Making Areas

Total Chairs

Runs

Fast quads

Area covered by snow makers (acres) distribution for resorts in market share

Total number of chairs distribution for resorts in market share

Big Mountain is very high up the league table of snow making area.

Big Mountain has amongst the highest number of total chairs, resorts with more appear to be outliers.

Summary:
Big Man currently charges $81.00. For a skier who stays an average of 5 days, the expected # of visitors is 350,000. In addition in a scenario ran where they add a run and increase the vertical drop by 150ft and installing an additional chair lift supports a ticket price increase of $1.25 which amounts to an expected $2181159.
However, the model does not take into account the additional capital expenditure and other on-going operations since data was not provided. Also finding that even increasing the longest run by .2 and adding 4 acres of snow made no difference in its prediction.

However there is still some further work that needs to be done. It appears that the model supports the higher premium ticket price Big mountain has. And that there is potential for it to be raised slightly more. But, seems to be some deficiencies that is linked to lack of information in certain areas such as facilities info and operating cost data and its comparison to the competitor resorts. Info on what facilities in the ski resort customers prefer to others. Gathering that data and comparing can help provide information on what facilities to add or change that can assist with raising ticket prices more. In addition, maybe hiring experts in the industry to test out the model and see if it's viable.