GUIDED CAPSTONE PROJECT REPORT FOR BIG MOUNTAIN RESORT

Data Science Methodology

We have worked through each step of the data science method to complete the capstone project

Step One: Problem Identification

After looking into the given information about the case study the problem has been identified and a problem statement is generated.

Problem Statement

How Big Mountain resort could alter its pricing strategy to lower the operation cost based on the 'best value to their customers' by comparing their facilities with other resorts in their market segment using data-driven strategies to recoup the increased operational cost of $1.54MM for installing new chair list this season, while keeping the profit margins at 9.2% and give an insight on annual revenue for the season over the next year.

Step Two: Data Wrangling

The second step of the data science method is data wrangling, also known as data munging or data cleaning. It involves several steps to ensure data quality and usability. Here are the steps we performed in data wrangling:

- We collected the relevant data from the ski resort data source.
- Explored the data to understand its structure and identified potential issues, such as missing values, outliers, and inconsistencies.
- Handled missing values, either by imputing them with reasonable estimates or removing them if they are too numerous or critical.
- Identified and removed any duplicate rows in the dataset.
- Examined outliers and decided whether they need to be corrected or removed.
- Converted data into a consistent format.
- Corrected any obvious errors or inconsistencies in the data.

Step Three: Exploratory Data Analysis

The step three of DSM is Exploratory Data Analysis. In this process we summarize our findings visualize results and analyze them to find any meaningful insights. The steps we performed in this process includes:

- Summarize the key statistics and insights derived from the EDA.
- Visualize the data distributions, correlations, and patterns.
- Identify any interesting trends or relationships in the data.

Step Four: Pre-processing and Training Data Development

This is the fourth step of the process; we pre-process and trained the data using different machine learning models and perform feature engineering to identify the best model which make appropriate predictions. This step we performed are:

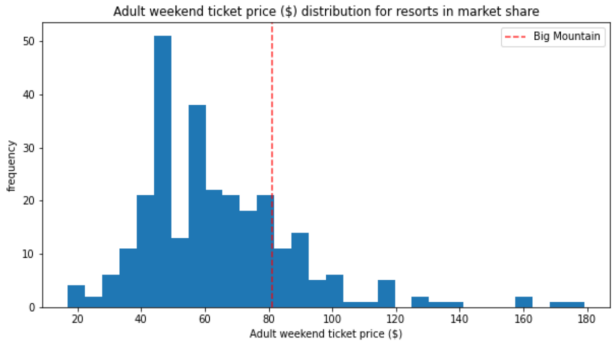
- Outlined the feature engineering techniques applied to create new features or transform existing ones.
- Normalized and scaled the data using different techniques.
- Used different machine learning algorithms to test and train the data.
- Created different models such as linear regression and random forest and compared the results.
- Provide the results and performance of each model.
- Trained the model using hyperparameter tuning.
- Created pipelines to identify the best approach and minimize errors.

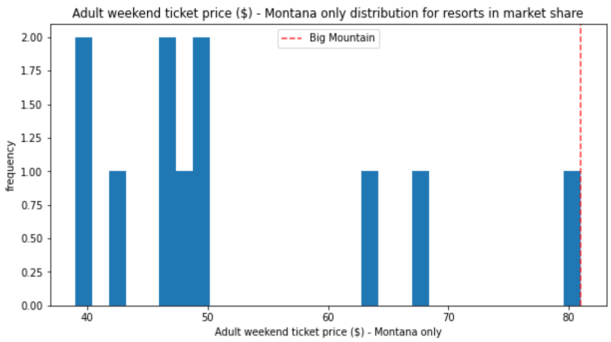
Step Five: Modeling

For the step five, modeling we performed the following steps:

- Chose the best model with highest accuracy and minimum errors.
- Documented the evaluation metrics chosen to assess model performance.
- Stated the best-performing model and its evaluation metrics.
- Described the scenario modeling using different scenarios for specific situations.
- Explained the reasoning behind choosing the winning model.
- Made suggestions for business executives on how to use the model in future.

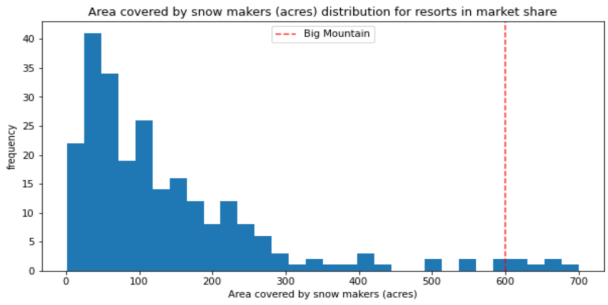
After performing all the necessary steps and choosing the best model for the optimized ticket pricing for The Big Mountain resort we further mapped and compared the ticket prices of Big Mountain with resorts in all other states and within Montana only.





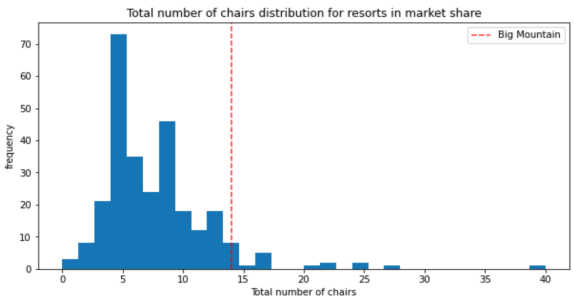
We then plotted compared different features of The Big Mountain resort with other resorts to see how well the Big Mountain is doing.

Snow making area



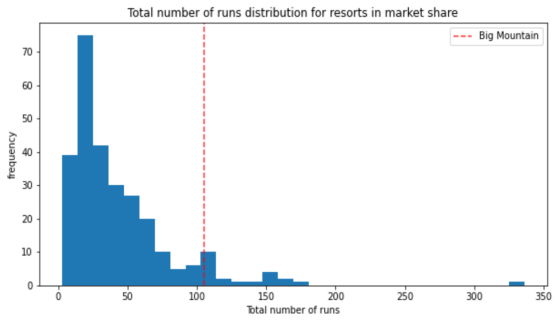
Big Mountain is very high up the league table of snow making area.

Total number of chairs



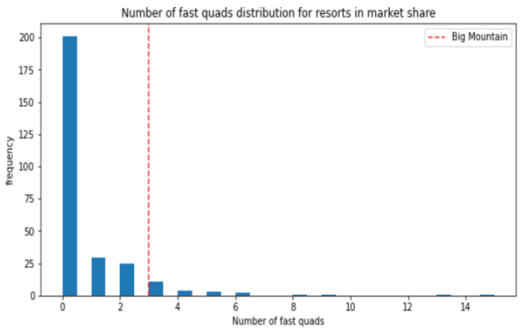
Big Mountain has amongst the highest number of total chairs, resorts with more appear to be outliers.

Runs



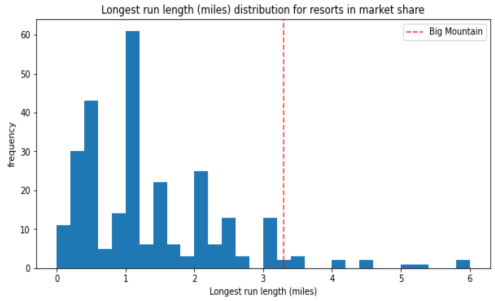
Big Mountain compares well for the number of runs. There are some resorts with more, but not many.

Fast quads



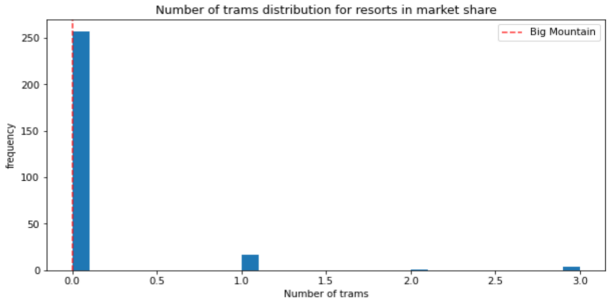
Most resorts have no fast quads. Big Mountain has 3, which puts it high up that league table. There are some values much higher, but they are rare.

Longest run



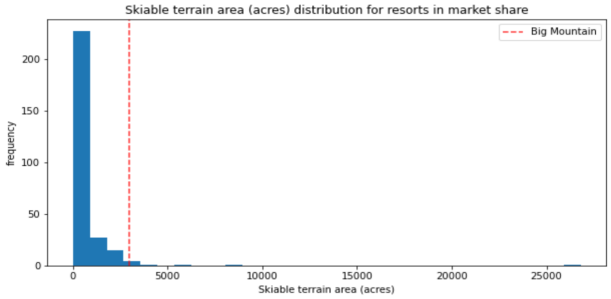
Big Mountain has one of the longest runs. Although it is just over half the length of the longest, the longer ones are rare.

Trams



The vast majority of resorts, such as Big Mountain, have no trams.

Skiable terrain area



Big Mountain is amongst the resorts with the largest amount of skiable terrain.

Different scenarios for the business-selected options were applied such that the expected number of visitors over the season is 350,000 and, on average, visitors ski for five days.

Scenario 1

Close up to 10 of the least used runs. The number of runs is the only parameter varying.

- For this scenario we created two plots, side by side, for the predicted ticket price change for each. If each of the expected visitors buys 5 tickets.

The results say closing 1 and 2 run makes no difference. Closing 3 successively reduces support for ticket price and so revenue. If Big Mountain closes 3 runs, it seems they may as well close 4 or 5 as there's no further loss in ticket price. Increasing the closures down to 6 or more leads to a large drop.

Scenario 2

Big Mountain is adding a run, increasing the vertical drop by 150 feet, and installing an additional chair lift.

- We predicted an increase in price by calling predict increase with a list of the features 'Runs', 'vertical_drop', and 'total_chairs'.

This scenario increases support for ticket price by $0.54 over the season.

Scenario 3

Repeating the previous one but adding 2 acres of snow making.

- Repeated scenario 2 conditions, adding an increase of 2 to Snow Making_ac.

The results were same as scenario 2 hence such a small increase in the snow making area makes no difference!

Scenario 4

This scenario calls for increasing the longest run by .2 miles and guaranteeing its snow coverage by adding 4 acres of snow making capability.

- In this scenario we predicted the increase from adding 0.2 miles to LongestRun_mi and 4 to Snow Making_ac.

The answer is : 0.20370370370369528. Even smaller increase. Although the longest run feature was used in the linear model, the random forest model (the one we chose because of its better performance) only has longest runway down in the feature importance list.

conclusion

In conclusion, the modeling results provide valuable insights for Big Mountain's business leadership to optimize ticket pricing, consider the additional operating cost of the new chair lift, and explore potential scenarios for future improvements. With data-driven recommendations and careful testing, the business can make informed decisions to enhance its competitiveness and guest satisfaction while maintaining financial viability.

Suggestions

Testing and progressing with run closures requires a technical approach to ensure minimal impact on guest experience and revenue while obtaining valuable data to inform decisions. Here are some tips for the resort so that they can make well-informed decisions regarding run closures:

- Data Analysis and Prioritization
- Small-Scale Testing
- Communicate with Guests
- Measure and Evaluate
- Operational Considerations
- Feedback Analysis
- Refine and Expand Testing
- Long-Term Impact Assessment
- Monitor and Adjust

Future Improvements

For future improvements some changes can be made to scenario 3 and 4 by adding more than 2 acres of snow making area and increasing the longest run by adding miles to see if it can help increase the ticket pricing for Big Mountain resort.