Guided Capstone Project Report

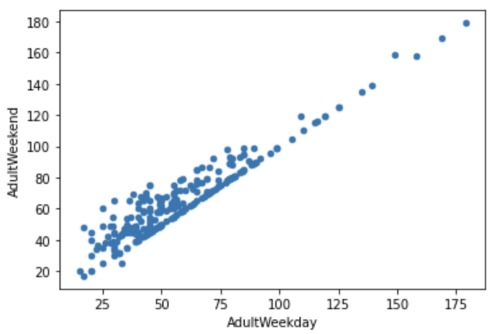
Problem Introduction

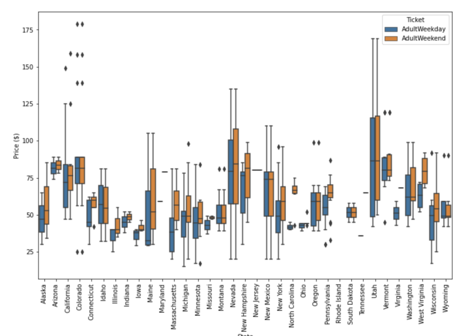
Big Mountain Resort is a ski resort in Montana with 350,000 visitors a year, looking to improve ticket pricing and revenue by capitalizing on its facilities. We are focused on identifying the facilities that Big Mountain Resort can invest in to support an increase in ticket price to drive higher revenue and areas the resort can divest from to lower operating costs. Our model aims to show which facilities have the greatest effect on ticket pricing and changes that could raise revenue if implemented.

Dataset, Features, and Preprocessing

The dataset of the ski resort market share that includes Big Mountain Resort was provided by Database Manager, Alesha Eisen. The dataset had 330 rows, each row a unique resort, and 27 columns representing different resort features. At the end of preprocessing, data had 277 rows and 25 columns.

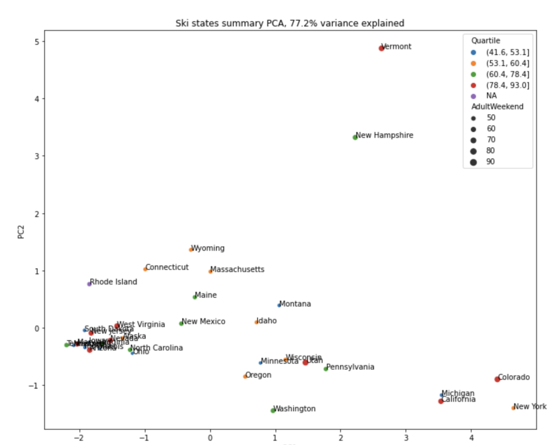
To choose the price target feature, we computed the relationship between ticket prices. A clear linear relationship was shown between AdultWeekday and AdultWeekend prices. The data had fewer missing AdultWeekend values than AdultWeekday values, as a result, AdultWeekend ticket price was chosen as the target feature for modeling.



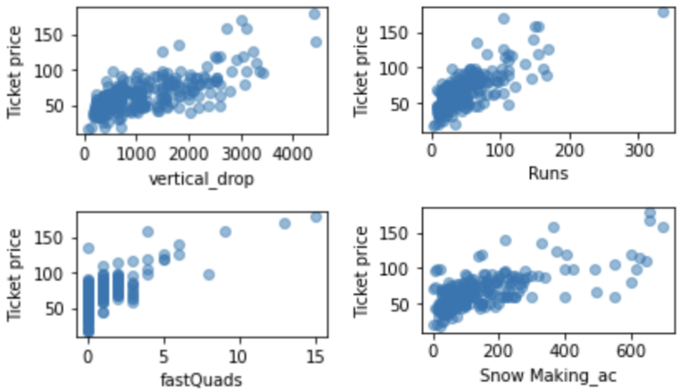


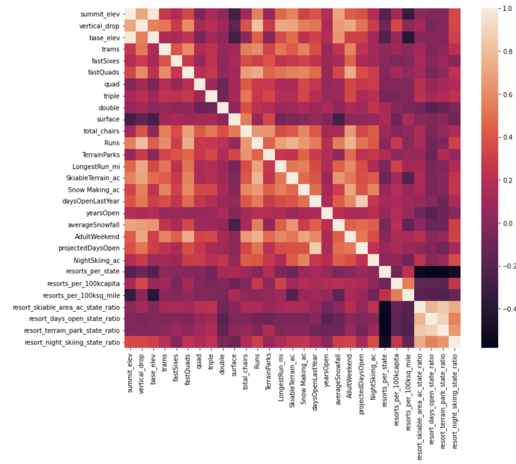
To decide whether to use all market share state data or a smaller subset, we investigated the relationships between our data and state by generating state total summary statistics and using population and land area data to create resort density ratios of resorts per 100k capita and resorts per 100k sq miles.

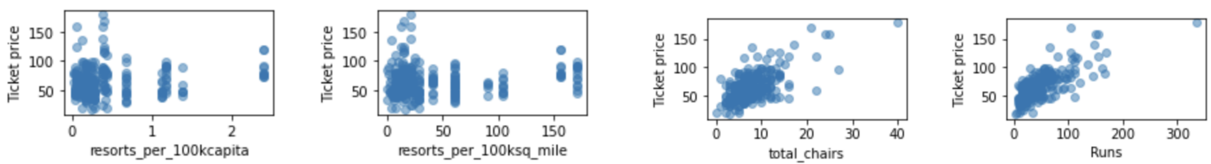
We used principal component analysis (PCA) to assess all feature variance by state. Though the ski states summary PCA figure showed that the first two components explained 77.2% of the variance seen, an obvious pattern did not emerge. This justified treating each state equally rather than focusing on a smaller subset, or only Montana. Some of the variability for the outlying states, like Vermont and New Hampshire, was explained by large values for resorts per 100k capita and resorts per 100k sq miles.



To quantify the competitiveness of each resort in their state, we showed the proportion of the state total metric attributed to each resort in that state and assessed correlations of all features to ticket price with a heatmap and scatter plots.



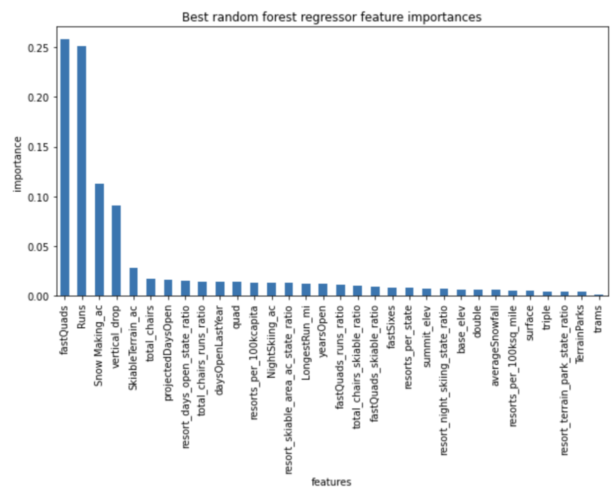




Models and Techniques

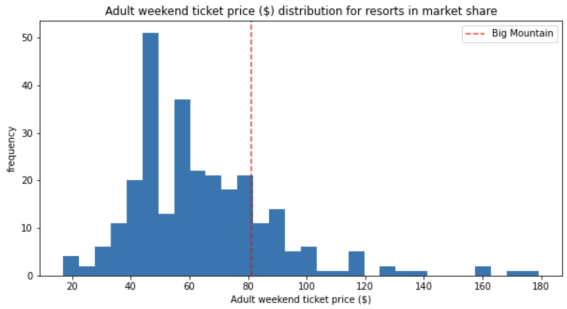
We gained a baseline of performance without modeling, predicting with average prices. The result had a mean absolute error of about $19. We then created a linear regression model and a random forest model to predict ticket prices and checked their efficacy with cross validation. Performance of both models on the test set confirmed their cross validation results. Both models showed the most important features as vertical drop, fastQuads, total chairs, runs, and snow making area, as shown below.

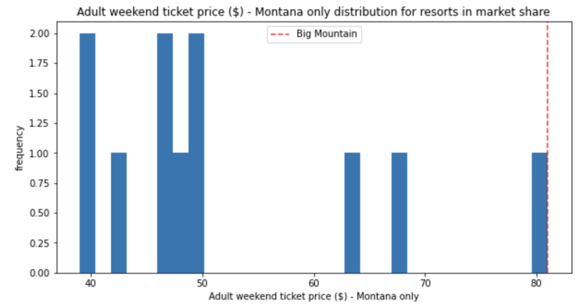
We chose to use the random forest model, as it had the lower cross validation mean absolute error by nearly $1 and exhibited lower variability. Improvement in the model cross-validation scores leveled off around 40-50, so Big Mountain Resort has enough data.



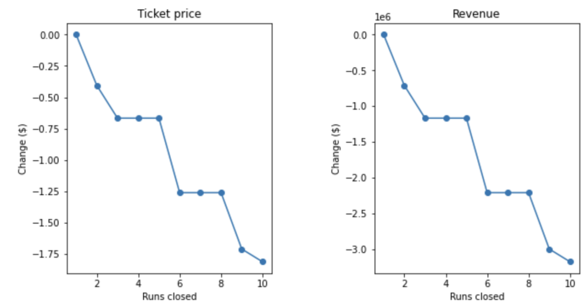
Results and Discussion

Big Mountain Resort's modeled ticket price is $95.87, and the actual price is $81.00. Even with the mean absolute error of $10.39, this suggests there is room for an increase. All Montana resort prices are under $100, with equal prices for both types of tickets. Our data showed resort prices over $100 have one price, and resorts below $100 often have price differentials. Big Mountain Resort could be poised to increase their weekend pricing or raise the price of both tickets according to the model and ticket price relationships; it should be noted that while Big Mountain Resort is in the middle of market share ticket price distribution, the resort is the highest priced in Montana.





We modeled some shortlisted scenarios from the business executives. Closing one run had no effect on ticket price, as shown in the figure below. Adjusting the snow making area and increasing the longest run did not support an increase in ticket price. Increasing the vertical drop by 150 feet, which requires the addition of a run and a chair lift, supported an increase in ticket price of $1.99.



The top most impactful features on ticket price, according to our chosen random forest regressor model, are fast quads, runs, snow making area, and vertical drop. Other important features were total chairs, longest run, trams, and skiable terrain area. Big Mountain Resort was in the upper range of distributions of every feature but trams, which a majority of resorts also do not have. Notably, Big Mountain Resort was also responsible for over 85% of nightsking in Montana.

Conclusion/Future

The model is based on the assumption that other resorts are not underpricing or overpricing; an assumption partially verifiable by modeling the expected prices of other resorts and comparing the results to reality. Competition between resorts in some areas that led to higher pricing could artificially inflate the model. Likewise, more resorts per capita in a large and less densely populated area could lead to competition. Big Mountain Resort's pricing is at the higher end for Montana, but middle-ground for all states. This could indicate a missing variable that lowers prices in Montana compared to other states, but verification would require modeling the other Montana resort prices to see if they have similar price discrepancies.

The business executives could have an understanding as to why their resort is priced so much lower than the modeled expectation. They should be consulted to gather any background or missing data points that could explain why Big Mountain Resort may be underpriced, despite its position in the upper range of nearly all most price-influential features. Additional data that could be useful are operating costs, modeled weekday prices as the target feature, and other average costs visiting skiers are responsible for such as food, lodging, and flights. A more remote resort may require a pricier flight and accomodations, necessitating lower ticket prices.

The data suggests a few recommendations. Big Mountain should maintain its fast quads. While the data does not show that expanding this feature further would be beneficial, because few resorts have more fast quads than Big Mountain Resort, the feature is important to ticket pricing and could help explain the discrepancy between the modeled and current Big Mountain Resort prices. The modeled scenario, increasing vertical drop by 150 feet, and adding a run and chair lift to support it, supports a ticket price increase of $1.99. Over the season, this could be expected to amount to $3,474,638 before inclusion of increased operating cost. Increased operating cost could be mitigated by closing another run elsewhere in the park, as closing a run was shown to have no effect on ticket price.

References

Springboard_Curriculum, DataScienceGuidedCapstone, (2021), GitHub repository, https://github.com/springboard-curriculum/DataScienceGuidedCapstone