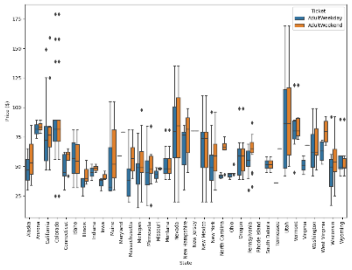
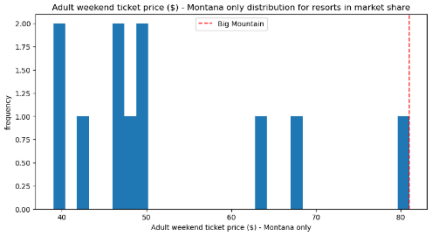
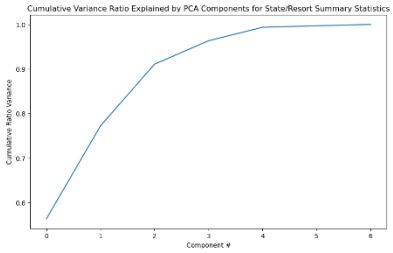
Big Mountain Resort, which services around 350,000, guests per year is looking for a model to base their pricing strategy around. The pricing strategy needs to be flexible enough to account for changes to facilities and quality while staying competitive with the broader resort market. To accomplish this the company gave us access to some of their internal data on other resorts and their own resort. The



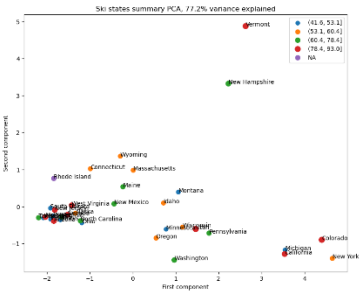


data provided showed how the resort ticket prices were high for their state, Montana, but not especially high across the country. The data did need to be cleaned up, with some columns being removed and others added for context. We checked for duplicates, which there were none, and for wrong or missing data. The last thing we did for the data wrangling phase was to find state summary statistics. Those statistics included population data, total area, and how the resorts are divided up among those states.

The exploratory data analysis started off with looking at how the data is grouped and skewed. Doing so showed that there is structure to the data. To analyze the data we ran it through principle component analysis to look at the how the variance of the data



relates to the different components. We learned there are a few components that account for most of the variance in the data, greater than 75% of it. Looking further in the data we see some counterintuitive data,



such as how total chairs does not relate to ticket price. Having more chairs does not seem to correlate to higher ticket prices. With the data we have available to us we can only speculate if that trend is because more chairs allow the resort to service more customers to take advantage of economies of scale.

For the processing phase we started off merely comparing the mean as a predictor. What we learned is the price varies greatly from the average price, which could indicate competitive pricing models from other resorts. From there we went to a basic linear regression model that predicted pricing much more closely. After tuning we found that 8 that most affected price and allowed for predictions in the linear regression model. What we moved to from there was a Random Forest model that performed even better than the linear regression model did after some hyperparameter tuning was performed. The Random Forest model narrowed down to four features that most predicted price. We moved forward with the Random Forest model to the next phase of the project after seeing it's performance over the previous models.

vertical_drop	10.767857
Snow Making_ac	6.290074
total_chairs	5.794156
fastQuads	5.745626
Runs	5.370555
LongestRun_mi	0.181814
trams	-4.142024
SkiableTerrain_ac	-5.249780

As things stand currently for Big Mountain Resort the model predicts that there is room for the resort to charge more than it currently is without the need for any changes to facilities. The current price the resort is charging is $81 for adult weekend tickets. The model suggests that the resort could charge as much as $95.87 with an error estimation of about $10. Even in the worst-case scenario for the error estimation the resort could be charging a few dollars more than it currently is. The resort presented a

handful of scenarios for investigation. The first scenario involved closing up to 10 runs in an effort to reduce costs. Our model predicts that reducing the number of runs by 3-5 has the least impact on pricing relative to the number of runs lost. The business will need to decide if the reduced income will be made up for by the reduction in costs associated with the runs shut down. The second scenario that was modeled involved adding a run, increasing the vertical drop by 150 feet, and installing an additional chair lift. Our model predicted that this would add $8.61 to the support for ticket price. Scenario three included the changes in the second scenario but also adding 2 acres of snow making. The resulting price prediction only added $1.29 to the support for ticket price. There was a fourth model proposed that would increase the longest run by .2 miles and adding 4 acres of snow making. Our model shows that would add no support to an increased ticket price. Given the predictions from the model the resort will need to decide which scenario is most cost effective taking into account the increased maintenance costs.

This model could be expanded with more data, including data on the costs of each feature and how those features attract and retain visitors. Learning more about how the business and industry operates would also allow for more precise modeling or for better understanding of how some of the features work together.