Big Mountain Resort recently installed a chair lift to increase visitors to the mountain, this had increased operating costs by $1.5 million. The pricing strategy originally was to charge a premium. Instead, they hired this data science team to implement a data-driven business strategy. The strategy below will be listed and documented to give Big Mountain Resort the best strategy to cover the costs.

**Problem Statement**

How can Big Mountain Resort select a better value for their ticket price within the next year that can also help cut costs by 10% without reducing the ticket price within the next year or support a premium price 5% above the market, also within the next year?

**Data Wrangling**

We were given a csv file that had 330 rows and 27 columns, and the wrangling consisted of cleaning the data, and checking the distributions of each data set. First, we checked to see if there were any missing data. FastEight missed around 50.3% of data, following NightSkiing\_ac missed around 43.3%. The AdultWeekday and AdultWeekend ticket prices had 3% missing one value, and 14% missing on both values. The first chart below shows what values are missing data.

Table

Description automatically generated

Looking at the distributes, we can look at fastQuads, fastSixes, and trams. These lack much variance away from 0 and and have a small number of relatively extreme values.

Timeline

Description automatically generated

**Exploratory Data Analysis**

After exploring the data, we have analyzed that there were multiple features such as: skiable\_area\_ac\_state\_ratio, resort\_days\_open\_state\_ratio, resort\_night\_skiing\_state\_ratio,total\_chairs\_runs\_ratio, total\_chairs\_skiable\_ratio, fastQuads\_runs\_ratio, fastQuads\_skiable\_ratio, Runs\_skiable\_ratio. There were also multiple categorical features such as: resorts\_per\_state, state\_total\_skiable\_area\_ac, state\_total\_days\_open\_last\_year, state\_total\_night\_skiing\_ac, resorts\_per\_100kcapita, resorts\_per\_100ksq\_mile.

Chart

Description automatically generated

After reading the graphs below, there is no relationship between state and ticket price. After reading this, we decided that the main features would be runs, total chairs, and fastquads because there is a strong correlation.

Chart, scatter chart

Description automatically generated

A picture containing calendar

Description automatically generated

**Model Preprocessing with feature engineering**

We developed a model by examining performance through taking the mean of the price. It did help us get an understanding of the price, but we decided to test through a linear model and the random forest model which ended up becoming more accurate. Taking the mean would be a little over $19.

Table

Description automatically generated

The linear model was off by $9, where we found the missing values using the median and mean values. Although this was better, the model still needed to be adjusted by adjusting the number of features. Through cross-validation, the value was set to eight features to focus on, vertical\_drop, Snow Making\_ac, total\_chairs, fastQuads, Runs, LongestRun\_mi, trams, and SkiableTerrain\_ac.

A picture containing table

Description automatically generated

Finally, the random forest model, like the linear model the missing values were found with the median and mean. The model revealed the top four features: fastQuads, Runs, SnowMakin\_ac, and vertical\_drop.

Chart, histogram

Description automatically generated

This became the winning model, as it produced the lowest margin of error and was consistent through the cross-validation results.

**Pricing Recommendation**

Blue Mountain currently charges $81, the modeled price suggests for the resort to charge $94.22. Our model does not support closing facilities, if we just increased the vertical drop by 150, then this would increase the ticket price by 10%, which would give us an increase of over $15 million.

**Conclusion**

In conclusion, after analysis, the best model to use would be the random forest model with charging $94.22. Running this price, this answers our initial problem statement and by raising the price, this would give Blue Mountain Resort more than enough money on covering the operating cost and charging an above market price on tickets.

**Future scope of work**

We are missing crucial pieces of data, such as off-peak hours data. This data could be missing data during the offseason, which could also have an increase of revenue. Employee costs are another avenue worth taking, if cutting employee salaries would help gain more revenue. The most important data to investigate would be to see how much it costs to maintain and do a run. These costs could directly indicate how much revenue Blue Mountain Resort can generate.