Rationale for this Project

Big Mountain Resort is a major ski resort of Montana, but we struggle to understand our exact value to the customers, which impacts our pricing strategy. We have been charging a simple premium over industry average pricing, but want to be more data-driven to capture additional potential revenue (10% increase in ticket price). We'd like to eliminate services that are less valued by the customer in a way that doesn't reduce our revenue also.

Could we generate 7~8 figure additional revenue per year?

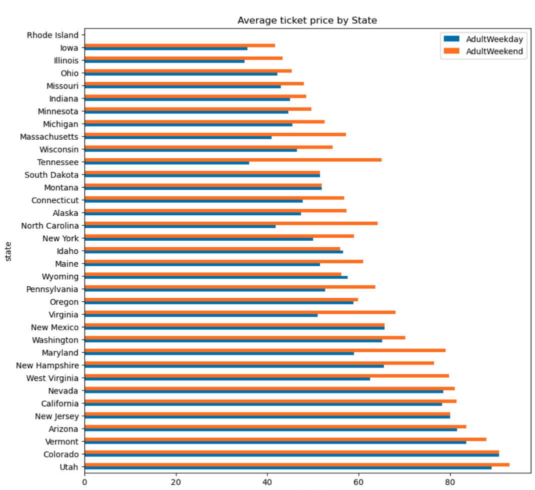
Available Data

We have a cleaned data set of 227 ski resorts from the US with the following information as columns (see right), and also acquired a data set containing US states and their population and landmass size from Wikipedia (link) as we anticipate relationships between states' size and population and ski resort ticket prices.

```
Column
------
Name
Region
state
summit_elev
vertical_drop
base_elev
trams
fastSixes
fastQuads
quad
triple
double
surface
total_chairs
Runs
TerrainParks
LongestRun_mi
SkiableTerrain_ac
Snow Making_ac
daysOpenLastYear
yearsOpen
averageSnowfall
AdultWeekend
projectedDaysOpen
NightSkiing_ac
```

By merging the two data sets and calculating ratios between certain features to capture the total supply of various ski resorts' offerings per state, such as # of resorts per state, total skiable area in a state, total skiing days per state, total # of terrain parks per state, night skiing offered per state, state's population and state's landmass. Think of these as additional columns added to the data set.

State Level Signal

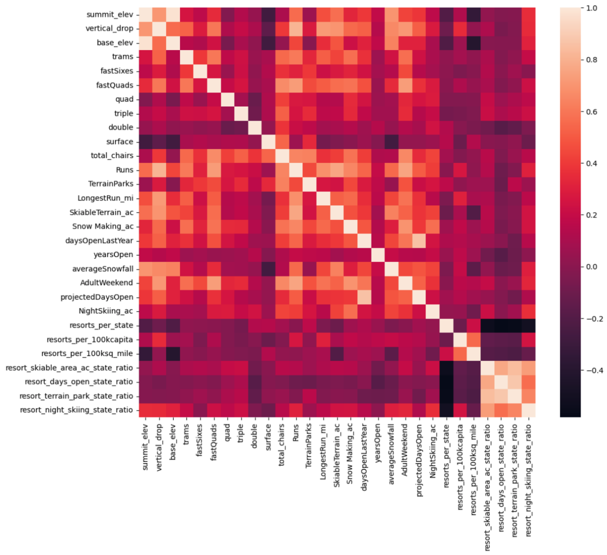


It'd be foolish to ignore the state variable and model our price purely based on the facilities and runs that we have. Big Mountain Resort in Montana can't expect to charge the same as an identical resort that's 20 minutes from Manhattan.

Using our intuition, we went to create several new data features, which were # of resorts per state, state total skiable area, state total days open, state total terrain parks, state total night skiing ac, state population, state areas sq mile, resorts per 100k population per state, resorts per 100k sq mile per state.
As seen in the chart above, Montana ranks below national median for its average ski resort ticket price. However, when we rank the states by various data features we have about them, it's impossible to find patterns that explain the difference in ticket price, which makes it an ideal candidate for Machine Learning and statistical modeling.

Resort Level Signal

At the individual resort level, excluding state information, we could map out the correlation between all the data features that we have about each resort using a heatmap.

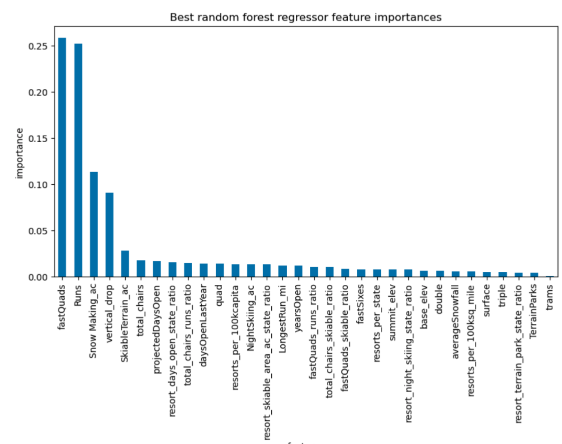


FastQuads, Runs, Snow Making_ac, total_chairs showed high correlation to weekend price, which seemed intuitive. resort_night_skiing_state_ratio was a less obvious signal that we discovered, which suggested that resorts could command a premium if they could dominate the night skiing supply in their states.

Both our state and resort data show some potential helping us model for our ticket price.

Machine Learning

After trying multiple versions of 2 different algorithms, we arrived at a version of Random Forest Model that provided best accuracy. The chart below shows the features by their importance in determining the accurate price. We are confident in this, because another ML approach (Linear Regression) also found the same top features to be important as well.



Ultimately a version of Random Forest Model proved most accurate for our use case of predicting Weekend Price, with Mean Absolute Error of 9.5, meaning that the price that our model predicts will be within $9.5 of the actual price.

Big Mountain Resort Price Recommendation

The winning Random Forest Model calculated that a price of **$95.87** can be supported based on the data that we have. The expected mean absolute error was $10.39, the floor of our price should be $85.48. Considering that our actual price is $81.00, there is some room for price increase.

In fact Big Mountain Resort ranks very highly nationally for the 8 key features that the model found, which supports the price increase recommendation.

- vertical_drop
- Snow Making_ac
- total_chairs
- fastQuads

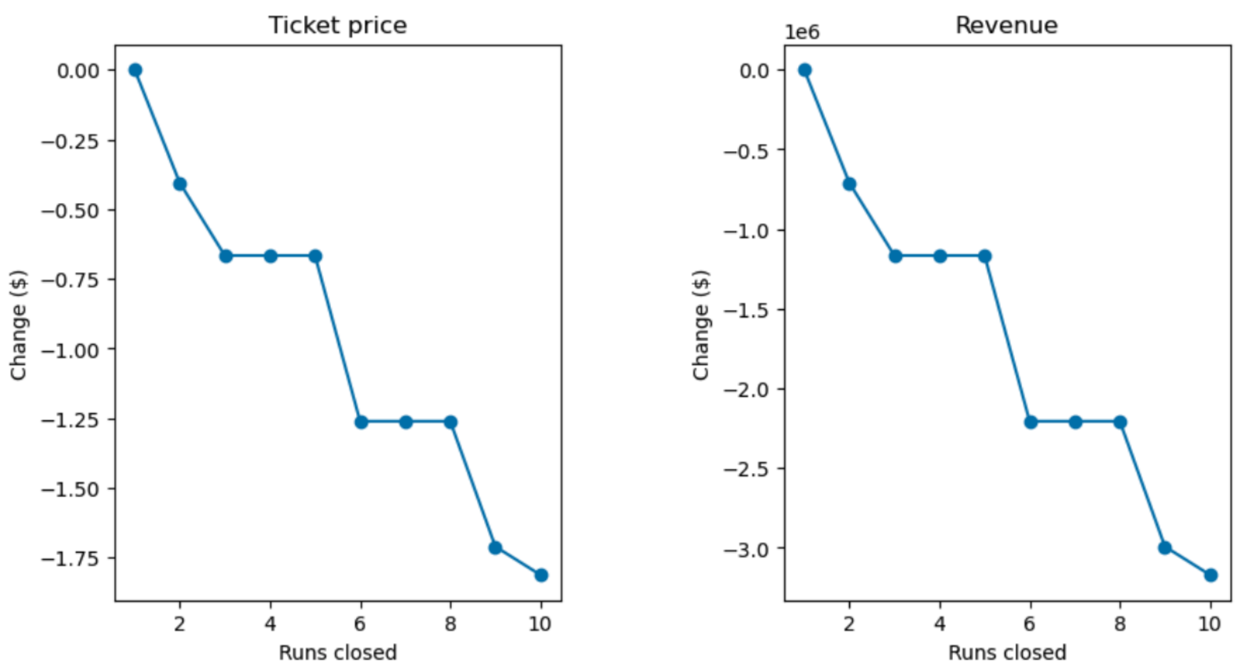
- Runs
- LongestRun_mi
- trams
- SkiableTerrain_ac

Big Mountain Resort Cost Reduction Scenarios

We further modeled 4 cost reduction measures provided by the business:

Assuming annual visitors of 350,000 purchasing 5 day tickets on average, here are the respective findings.

1. Permanently closing down up to 10 of the least used runs:
    a. No difference in revenue when closing 1 run, followed by incremental drops in steps as shown below.



2. Increase the vertical drop by adding a run to a point 150 feet lower down but requiring the installation of an additional chair lift to bring skiers back up, without additional snow making coverage
    a. This support ticket price increase by $8.61 and increase in season's revenue by $15,065,471.
3. Same as number 2, but adding 2 acres of snow making cover
    a. This supports ticket price increase by $9.9 and season's revenue by $17,322,717.
4. Increase the longest run by 0.2 mile to boast 3.5 miles length, requiring an additional snow making coverage of 4 acres
    a. The model expects no difference in price based on this scenario.

My recommendation after analyzing the 4 scenarios would be to assess the cost reduction impact of closing down 1 ~ 5 least used runs, which can help offset the cost of adding an additional chair lift to support a new run that adds 150ft of vertical drop.

For the run closures, I recommend starting by closing 1 run at a time to assess the impact of ticket sale for the first 3 closures, and if the results are in line with the models prediction, close 2 more at the same time.

Future Scope of Work

By collecting data around the cost of operating various facilities and assets, we should be able to more robust cost reduction and profit prediction models. This was a key data type that we were missing in this analysis.