**Big Mountain Ski Resort**

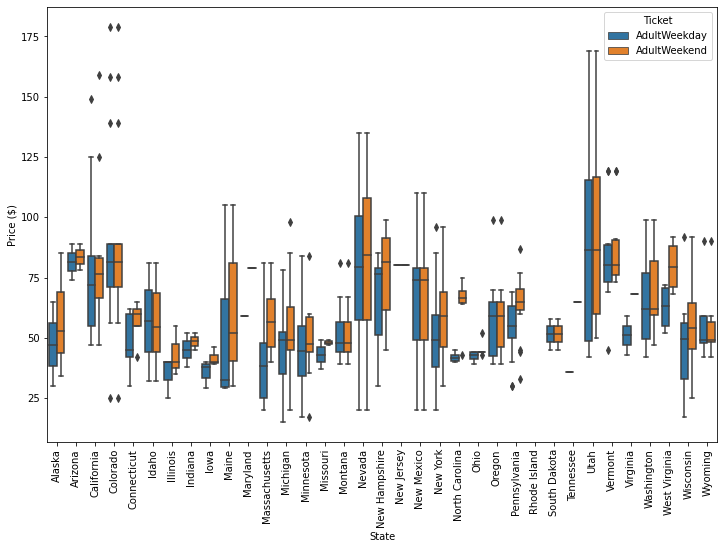
**Project Report**

**Problem statement**

Review ski resorts ticket prices and amenities across the country to come up with a pricing model for ski resort tickets in its market segment, and a model for future facility investment plans by next season.

**Data Wrangling**

After handling missing and wrong data points, we derive state-wide summary statistics for our market segment.



The plots shows Montana’s price is on the lower-midum range.

**Exploratory data Analysis**

We explored a state-wide summary dataset to understand what market segment we are in. We also scaled the data and did a PCA (principle components analysis) to figure out wheatear is a obvious pattern between states and ticket price.

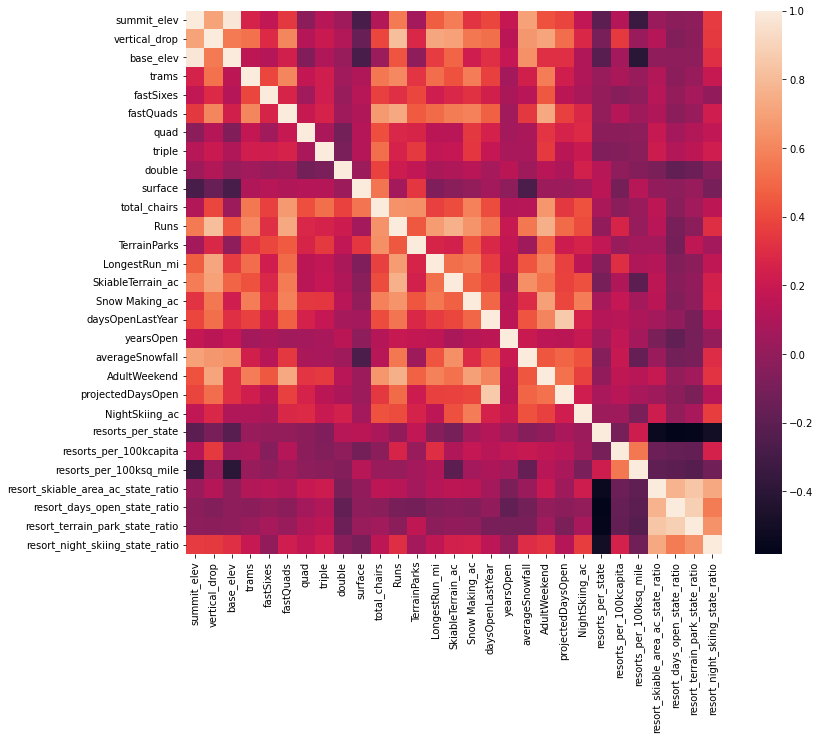
A picture containing chart

Description automatically generated

The red points representing the upper quartile of price can be seen to the left, the right, and up top. There's also a spread of the other quartiles as well. In this representation of the ski summaries for each state, which accounts for some 77% of the variance, there is not a clear pattern with price.

So we work towards building a pricing model that considers all states together, without treating any one particularly specially.

Then we merged state summary features into the ski resort data and added some intuitive features to explore the feature correlation.



From the heatmap, we can see the target feature, AdultWeekend ticket price, correlated with the following features: fastQuads ,Runs and Snow Making\_ac, resort\_night\_skiing\_state\_ratio.

**Model Preprocessing and model buildng**

We dropped Name, state, and Region columns for our model, so we are left with all numeric features. We split the data into training and test sets with the ratio of 70%:30%. Then we explored if mean is a good predictor, turns out it’s not. Then we tried linear regression, random forest; The random forest model has a lower cross-validation mean absolute error by almost $1. It also exhibits less variability. And we extracted the feature importance from the best model:

Chart, histogram

Description automatically generated

**Winning model and scenario modelling**

Then we refit the best performing random forest model on the whole dataset (excluding Big Mountain) , Big Mountain Resort modelled price is $95.87, actual price is $81.00. Even with the expected mean absolute error of $10.39, this suggests there is room for an increase.

Then we compared the most important features to other resorts, including vertical\_drop, Snow Making\_ac, total\_chairs, fastQuads, Runs, LongestRun\_mi, trams, SkiableTerrain\_ac, in order to see what we can to do reduce cost and increase revenue.

Chart, histogram

Description automatically generatedChart, bar chart

Description automatically generated

Chart, histogram

Description automatically generatedChart, histogram

Description automatically generated

A picture containing histogram

Description automatically generatedChart, histogram

Description automatically generated

Chart

Description automatically generatedA picture containing shape

Description automatically generated

Then we modeled the following scenarios:

1. Permanently closing down up to 10 of the least used runs. This doesn't impact any other resort statistics.
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
3. Same as number 2, but adding 2 acres of snow making cover
4. Increase the longest run by 0.2 mile to boast 3.5 miles length, requiring an additional snow making coverage of 4 acres

**Recommendation**

Ticket price

Current ticket price at Big Mountain is $81. Based on the model, Big Mountain could support a ticket price of $95.87 without making any changes to the current features. The expected mean absolute error is $10.39, but it still suggests there is room for an increase. This assumes the planned chairlift is installed and operations. So we recommend increase the price to $95.

Amenity changes

1. **Permanently close up to 6 of the least used runs.** Closing 1 run does not make a difference in ticket price/revenue. A small drop in ticket price/revenue is noted after closing 2 runs. Closing 3-5 yields the same drop of about **$0.70** in ticket price and **$1.20** in lost revenue. Chart, line chart

   Description automatically generated.
2. **Increase the vertical drop**.Based on our model, adding a run, increasing the vertical drop by 150 feet, and installing an additional chair lift increases support for ticket price by $8.61.
3. Do not add 2 acres of snow making or increasing the longest run, the model suggest a small and no increase respectively.

**future scope of work**

Additional information is required before implementing any of the above proposed feature changes to ensure the overall profit pays off. The biggest deficiency in the data is that we don't have any data on operating cost. To improve the model, the company need to source and incorporate operating cost data into the consider.