Big Mountain Resort

Report by George Luke

Context

This project aims to build a predictive model for ticket price based on a number of facilities, or properties, boasted by resorts (*at the resorts).*

This model will be used to provide guidance for Big Mountain's pricing and future facility investment plans.

Problem

Big Mountain suspects it may not be maximizing its returns, relative to its position in the market. It also does not have a strong sense of what facilities matter most to visitors, particularly which ones they're most likely to pay more for.

Recommendation | Key Findings

Presently, Big Mountain is charging about 81 dollars, but the model predicts 96, with a MAE of 10.24.

We have about 5 dollars of increased ticket price room, given the model. This could be as large as 10M in revenue we might be missing out on, given an assumption of 5 tickets purchased at a time.

Potential total revenue change is 10M, and that we might have the opportunity to decrease our relative operating costs, by making a few adjustments.

Since the models suggest closing 1 run might not result in much difference in ticket cost, we could experiment with dropping one and see how that impacts revenue.

Modeling Results : Model 1 - Linear Regerssion

We expected that we'd get a ticket price within $9 or so of the real price given MAE (better than the $19 of the not-model of just guessing off the mean), or around $12 (if we used MSE), using our linear regression model.

When we compared performance of the linear regression models on mean vs. median imputed values, the difference wasn't significant.

Cross validation allowed us to get an estimate of the variability of the performance estimate within 2 std's of the mean, at 47-86% of the data set being covered by our model.

We used the GridSearchCV function to help us narrow our k value down to 8 core features. This allowed us to see a variety of very interesting coefficients and features.

+'Vertical drop' is the biggest positive feature.

+Some other features were 'he area covered by snow making equipment, total chairs, fastQuads, the number of Runs, the longest Run.

+We also found some negative features: trams, and the area covered as skiable terrain. This last one was surprising to us. It's possible this is related to larger resorts allowing resorts to charge less per ticket, given more visitors hosted and our missing

Modeling Results: Model 2

-We then used a Random Forest model.

-After doing some model performance assessment via mean absolute error, we found that the random forest model had a lower cross-validation MAE by $1, with less variability. **This makes the random forest better than the linear regression model, and better than our not-model**.

Analysis

We dropped 14% of the rows without price data.

fastSixes and Trams are features that we had some mild concern over, but decided to leave for now.

We added features to the data to capture state-wide market size. We did this by aggregating

- TerrainParks

- SkiableTerrain_ac

- daysOpenLastYear

- NightSkiing_ac

We settled on our target value of weekend prices, that we'll set up our other variables to predict in our machine learning model.

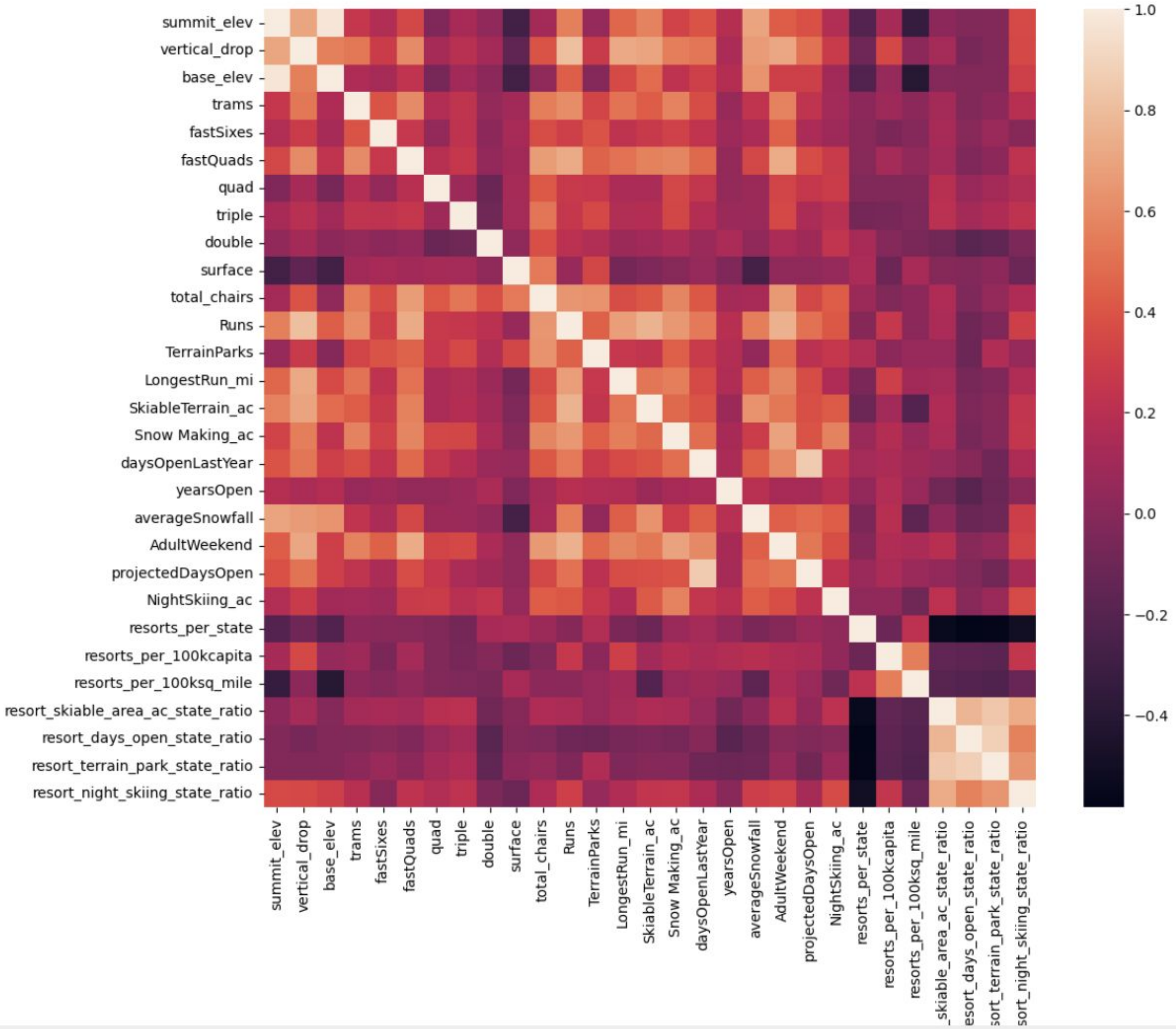
Analysis Pt 1

1. There's a strong positive correlation with vertical_drop.
2. fastQuads seems very useful.
3. Runs and total_chairs appear quite similar and also useful.
4. resorts_per_100kcapita shows something interesting that you don't see from just a headline correlation figure. When the value is low, there is quite a variability in ticket price, although it's capable of going quite high.
5. Ticket price may drop a little before then climbing upwards as the number of resorts per capita increases.

Analysis Pt 2

1. Ticket price could climb with the number of resorts serving a population because it indicates a popular area for skiing with plenty of demand. The lower ticket price when fewer resorts serve a population may similarly be because it's a less popular state for skiing.
2. The high price for some resorts when resorts are rare (relative to the population size) may indicate areas where a small number of resorts can benefit from a monopoly effect. It's not a clear picture, although we have some interesting signs.
3. The more chairs a resort has to move people around, relative to the number of runs, ticket price rapidly plummets and stays low. (What we may be seeing here is an **exclusive vs. mass market resort effect**; if you don't have so many chairs, you can charge more for your tickets, although with fewer chairs you're inevitably going to be able to serve fewer visitors)
4. Your price per visitor is high but your number of visitors may be low.
5. Something very useful that's missing from the data is the number of visitors per year
6. having no fast quads may limit the ticket price, but if your resort covers a wide area then getting a small number of fast quads may be beneficial to ticket price.

Our target feature is still AdultWeekend, which we set as "Total Price" above.


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
resorts_per_state
resorts_per_100kcapita
resorts_per_100ksq_mile
resort_skiable_area_ac_state_ratio
resort_days_open_state_ratio
resort_terrain_park_state_ratio
resort_night_skiing_state_ratio
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4