GUIDED CAPSTONE PROJECT REPORT

Summary of price modeling for Whitefish Ski Resort (formerly Big Mountain)

Deanna M. Church
2024-05-21

Executive Summary

This study was undertaken to investigate ski-resort features that drive value for customers. By identifying features that drive value I wished to understand areas in which Big Mountain could potentially cut cost, without impacting the customer experience or cutting ticket prices. Additionally, I wished to identify the areas where investment may have the biggest impact on revenue generation for the upcoming season.

The features that most consistently drive value for customers are number of fast quads, vertical drop, number of nuns and snow making ability.

The recommendations are as follows:

- Without making any changes, increase the ticket price by $9.
- Close 2-3 of the least used runs

If the above actions increase revenue without impacting the customer base, the following items would be the next priority:

- Close 2-4 more of the least used runs
- Increase the length of the vertical drop and install a new chair lift.

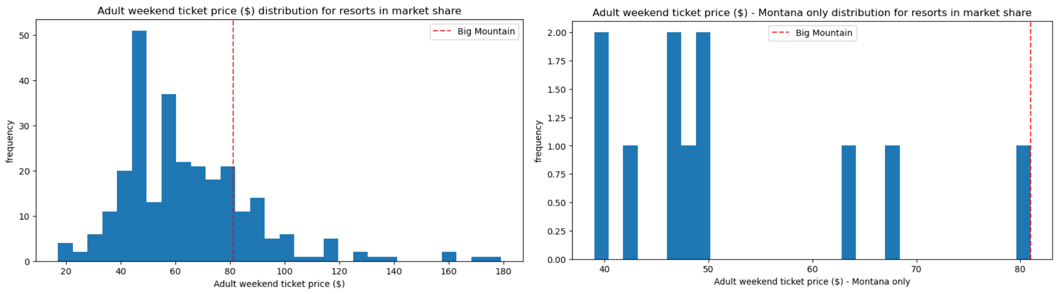
Recommendation Details

The recommendations are provided in an order that allows us to test the validity of the model. Initial recommendations are weighted toward low-cost items that are easy to reverse. As the model is validated, more aggressive changes can be implemented.

Big Mountain provides good value for the price, but is the highest price resort in Montana

> **Note**
>
> The Adult Weekend Price was used for this modeling exercise. The Adult Weekend and Adult Weekday prices are strongly correlated, and there was slightly more data for the Adult Weekend Price.

When evaluating ticket prices for the current amenities Big Mountain offers, the model identified $95.87 as the target price (compared to the current price of $81). The model as an average error of +/- $10.39.



(a) All ski resorts

(b) Montana ski resorts

Figure 1: Pricing information, with Big Mountain highlighted.

! Recommendation

Increase the ticket price by $9. There is little expense associated with this and it is easy to roll back. Note: It is unclear is local skiers are more price sensitive than out of state visitors- this demographic is worth monitoring.

Close 2-3 of the least used runs

While an increased number of runs is a revenue driver, the model suggests that closing a few of the least of the least used runs would have little negative impact on revenue.

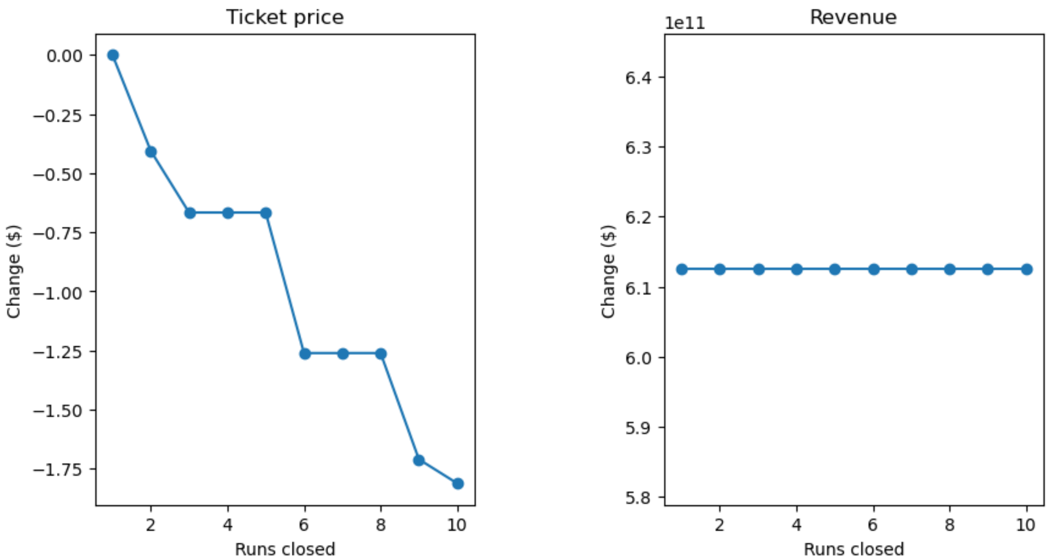


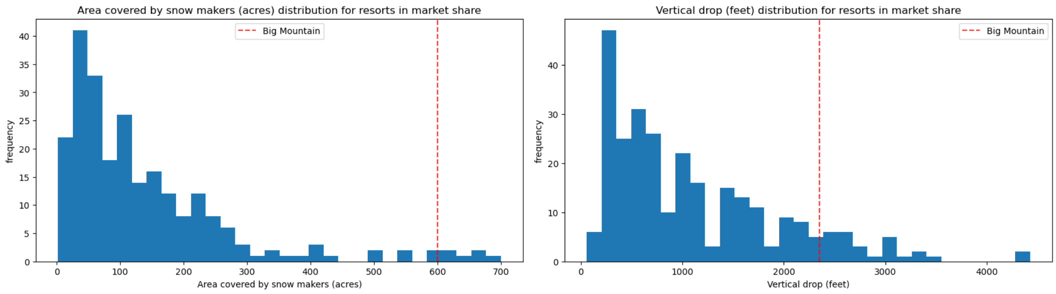
Figure 2: Run closures impacting ticket pricing and revenue

! Recommendation

Start by closing 2-3 runs and monitor the impact on customers. While closing more runs does have a negative impact on ticket prices, these are within the margin or error of the model, so it is unclear how meaningful this is. If the outcome is good, the closure of additional runs could be warranted.

Increase the length of the vertical drop and install a new chair lift.

Two scenarios were modeled with respect to vertical drop (with a new chairlift) with and without additional snowmaking capacity. Both scenarios pointed to the same ticket price and revenue impact: increase ticket price by $1.99. Over the season, this could be expected to amount to $3,474,638. Big Mountain already has tremendous snow making capacity, and there seems to be more room to grow by increasing the vertical drop.



(a) Ski making capacity

(b) Vertical Drop

Figure 3: Ranges of ski making capacity and vertical drop, with Big Mountain highlighted.

Future scope of work

There is additional information that may be useful for making stronger suggestions.

- the operating cost of each activity (snow making, adding a run, maintaining a run, etc)
- the annual number of visitors for each resort- this would have allowed us to dissect some of the complicated data around ski terrain, runs and chair numbers.
- the make up of the visitors (in-state vs. out-of-state) as they may exhibit different behaviors and have different costs sensitivities.

In terms of amenities, Big Mountain performs well compared to other resorts. Though, it is worth noting it is the most expensive resort in Montana. While the model suggests a higher price can be supported (an almost $15 increase)- the error on the model is +/- $10, so there is still a decent amount of uncertainty there. Leadership may be surprised that the model suggests they are providing more value than the ticket price would suggest but I would caution against upping ticket prices to the maximum initially, as all model building is a hypothesis generating activity, and it is worth while to think of ways to test the model to gain confidence in it.

If leadership find the model useful, I could develop an interactive application that allows leadership to explore changing features on their own. Additionally, the business could consider putting more resources into gather additional features on these ski resorts, or at least better understanding the demographics of its customers.

Conclusions

The model produced using the current data set provides some insights into amenities that drive value for customers. The mean error in the model is +/- $10, and many of the predictions fall within this error range. However, these do provide some testable hypotheses that can be evaluated. Many of these changes can be made at little or no cost, and are easily reversible. This provides a pathway for moving forward on increasing revenue and/or cutting costs.

Analysis Details

Initial Data Cleaning

The data were imported into a python dataframe. Missing and anomalous values where identified.

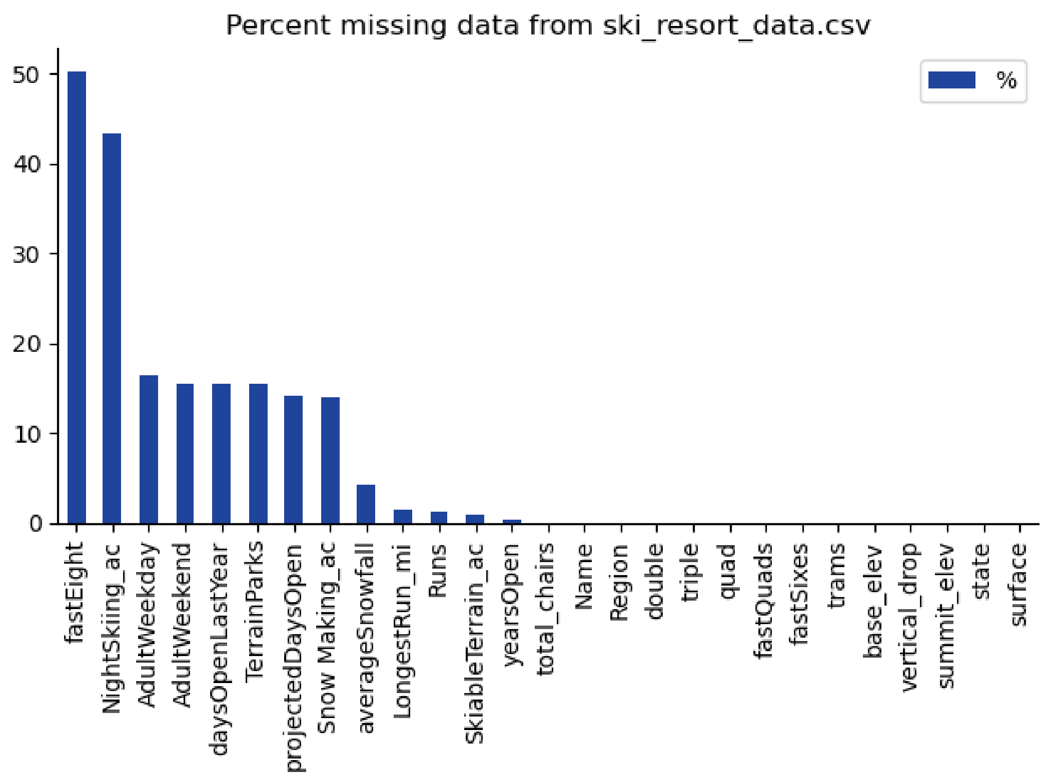


Figure 4: Percent missing data by data item

Outlier and missing data were handled as followed:

- Clear outlier in 'SkiableTerrain_ac' for Silverton Mountain. Investigated what the real value is and corrected this (from 26,819 -> 1819) in the main dataframe.
- Identified that 'Snow Making_ac' is likely wrong for Heavenly Mountain Resort. Did not fix as this resort is also missing pricing data.
- Dropped fastEight column as it was missing about half the data, and the rest of the data was heavily skewed to 0 (almost all).
- Pine Knob Ski Resort has an obvious error in 'yearsOpen'- dropped row due to difficulty of resolving accurately.
- Dropped all rows where both price values were missing.
- Dropped 'AdultWeekday' column.
- Dropped rows missing values in 'AdultWeekend' column.

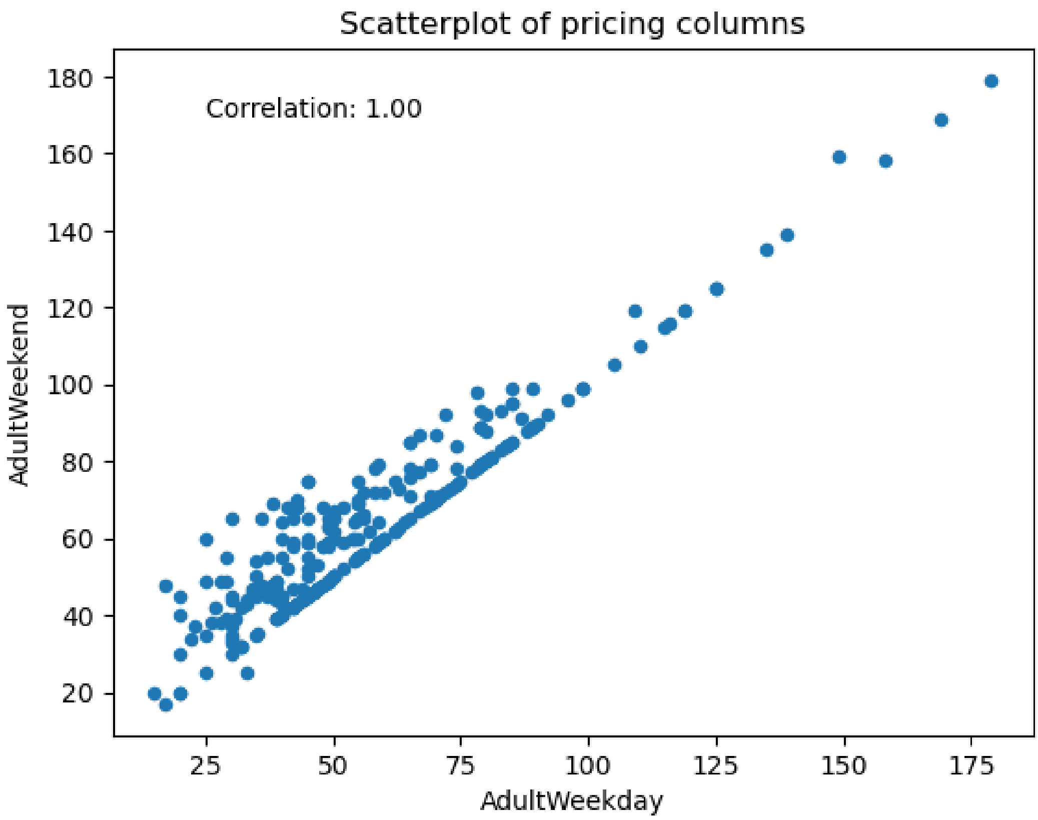


Figure 5: Scatterplot of Adult Weekend and Adult Weekday prices

The two prices are highly correlated, and there is slightly more data available for the Adult Weekday price, so this was chosen as the target features.

Additionally, some statewide information was imported and state summary fields calculated.

Exploratory Data Analysis

Exploratory analysis was performed to identify features that may be most valuable for building models. Initial work was performed to understand if state labels would be valuable. While the state labels themselves are not valuable, some aspects of the state information are useful for context, if not model building. Additionally I explored the relationships between the variables. Some of these are trivially correlated (such as summit elevation, base elevation and vertical drop). Many of these relationships are quite complicated.

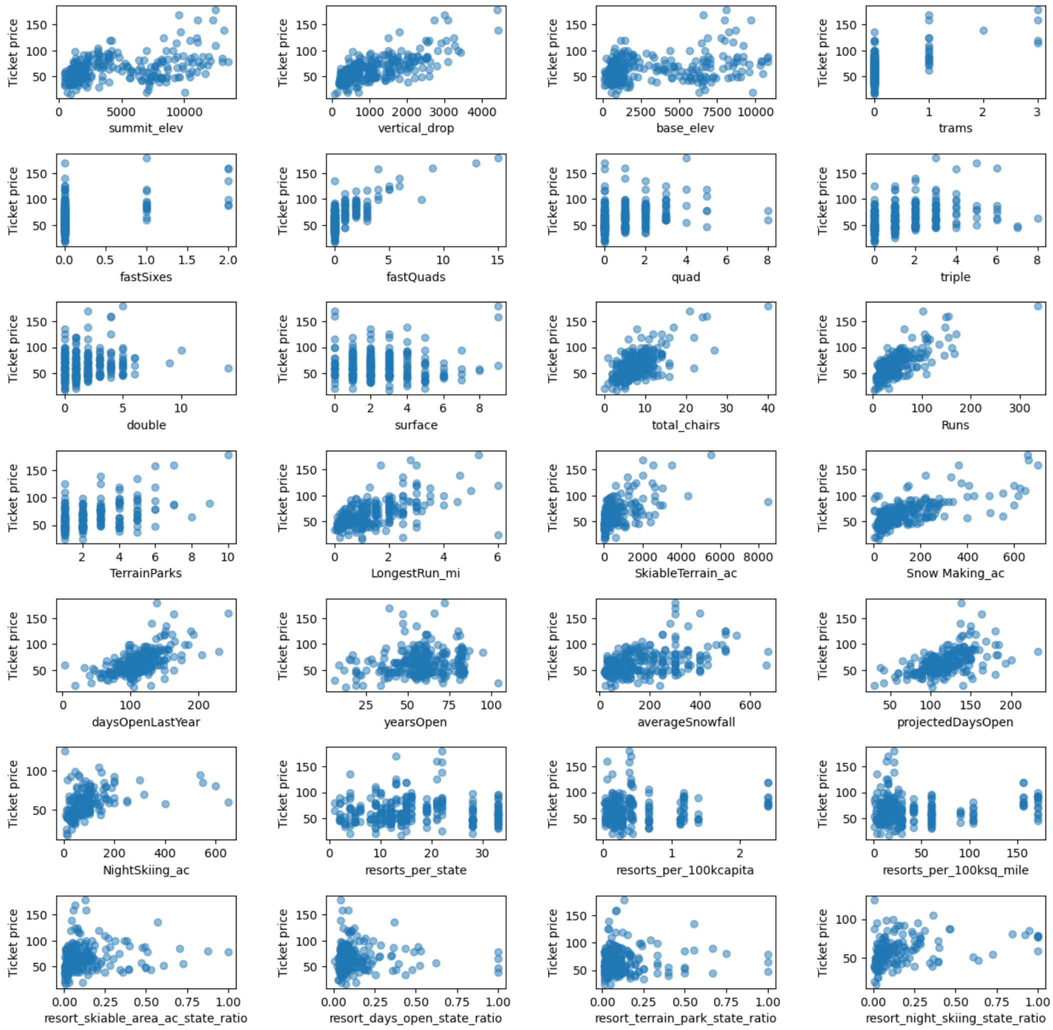


Figure 6: Feature correlations to ticket prices. Ticket price is on the y-axis and the feature of interest is on the x-axis

Model preprocessing and training

The information for Big Mountain was removed and the data was divided into a train (70%)- test (30%) split. I determined that just using the mean was not a good predictor for this data set, and moved on build models using the training data, with 5-fold cross validation (that is, we used different chunks of the training data to train the model multiple times and then averaged the results). Missing values were imputed using the median value for each column. Both Linear Regression and Random Forest models were generated. The performance of the Random Forest model was slightly better and this was chosen for the final modeling.

Modeling

The information for Big Mountain was removed, and the rest of the data was used to fit the Random Forest model. I think used this model to test various scenarios and assess the impact of these changes on ticket price.