Problem Statement Worksheet (Hypothesis Formation)

How can Big Mountain Resort support its increasing operational costs, $1,540,000 this season, by adjusting its operational facets such as snow cover, number of chairlifts, or ticket prices

H

1 Context

Big Mountain Resort is a ski resort located in Montana. Big Mountain Resort has recently installed an additional chair lift to help increase the distribution of visitors across the mountain which increases their operating costs by $1,540,000 this season. Traditionally Big Mountain has priced its tickets based on market price. Big Mountain wants to explore if data about its facilities can give some insight into pricing its tickets more suitably, or better yet make changes to its facilities to cover the higher operational costs. Some possible changes include adjusting its snow cover, number of chairlifts and/or ticket prices (or other factors)

2 Criteria for success

Costs for this season reduced by 1,540,000 or greater

3 Scope of solution space

- Available facilities and operations and ticket prices

4 Constraints within solution space

- **Lack of clearly defined success criteria – is the goal of the business to decrease the operational cost by the cost of the new chairlift?**
- **Lack of better usage data – for example number of users at each location, daily users, nightly users**

5 Stakeholders to provide key insight

- Big Mountain management

6 Key data sources

- Big Mountain operations data: Link

Recommendations

Pricing recommendations:

- Given that the cost of operating the additional chairlift is USD 1.54 million, we would still have to cover USD 1.19 M after increasing the ticket price. But this seems to be the best way to cover the additional costs

Strategy Recommendations:

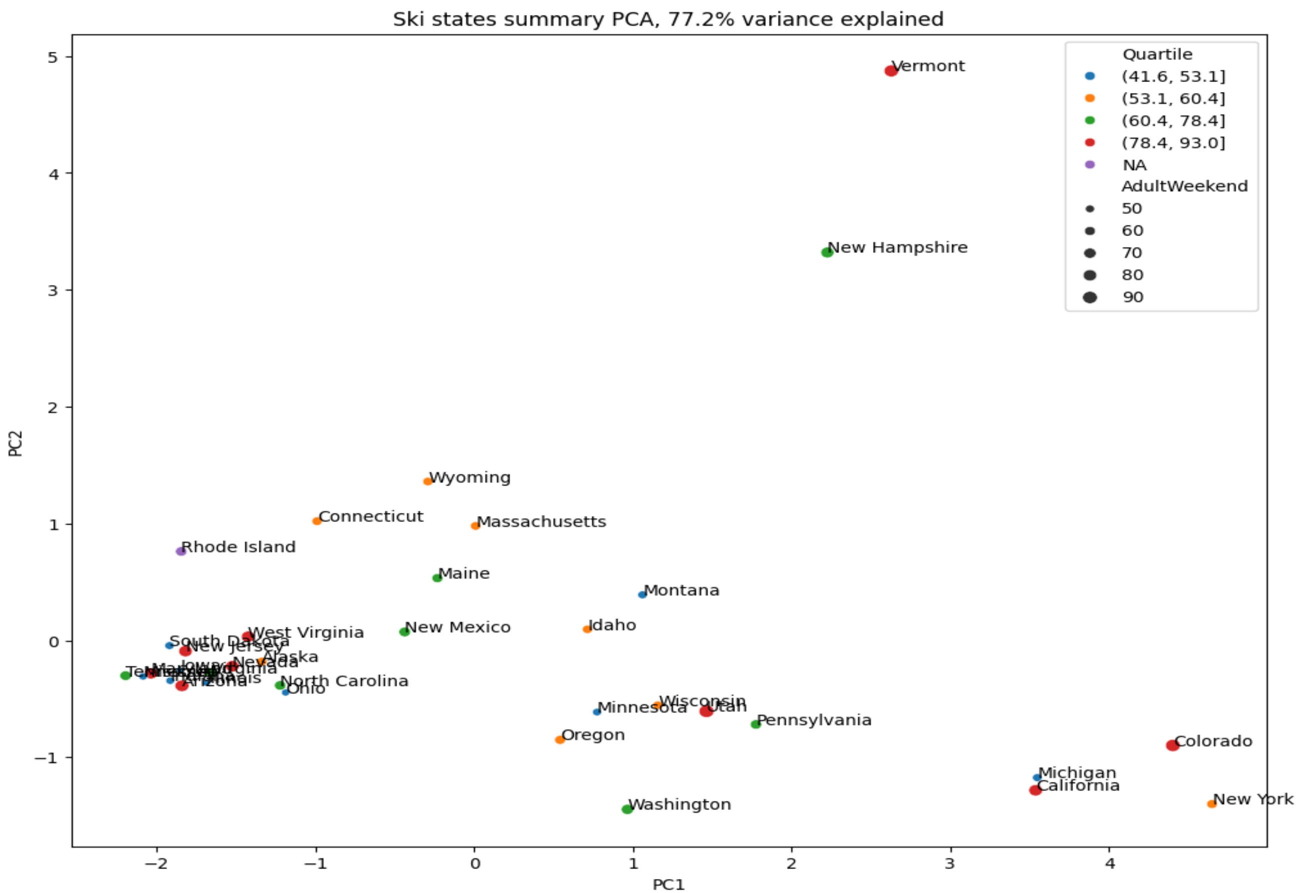
- We could temporarily partition the run into two parts, implementing a separate lift system for the additional 150 ft. We could entice users to agree for testing the additional drop using discounts or offers to other facilities. Could we go higher in price, upto $4 instead of just $2, based on our initial prediction of $96(-10 MAE)? When we test this additional price increase, does it discourage users?

Data recommendations:

- One very useful information would be the number of visitors for each resort. This would help us make better estimates for how features impact the buyers' attitudes. For example, would lower ticket prices by eliminating some runs attract more customers?

Feature Engineering

- After reviewing our ski data, the target variable is ‘AdultWeekend’ prices. This is because it had similar value to Weekday prices, while also having fewer missing values
- Initial PCA analysis also did not reveal a patter of price w.r.t. state, hence state isn’t by itself considered a factor in pricing tickets



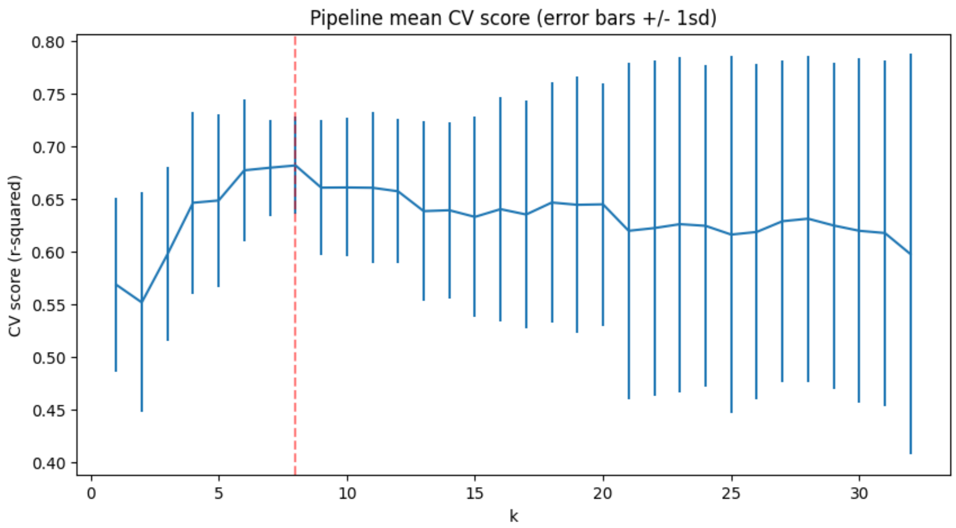
Modeling

Simple Average Predictor

- First a simple average was used to predict the ticket price. This produced a large error ($19) and as expected, very poor fit (R^2 = 0).

Linear Regressor Predictor

- Next a simple LinearRegressor model was tried. Missing values were imputed with median prices.
- This time, we got much better fits of R^2=0.72 and a MAE of $9.4, much less than the $19 from earlier
- Cross-validation tests to select the K most relevant features gave us a K=8



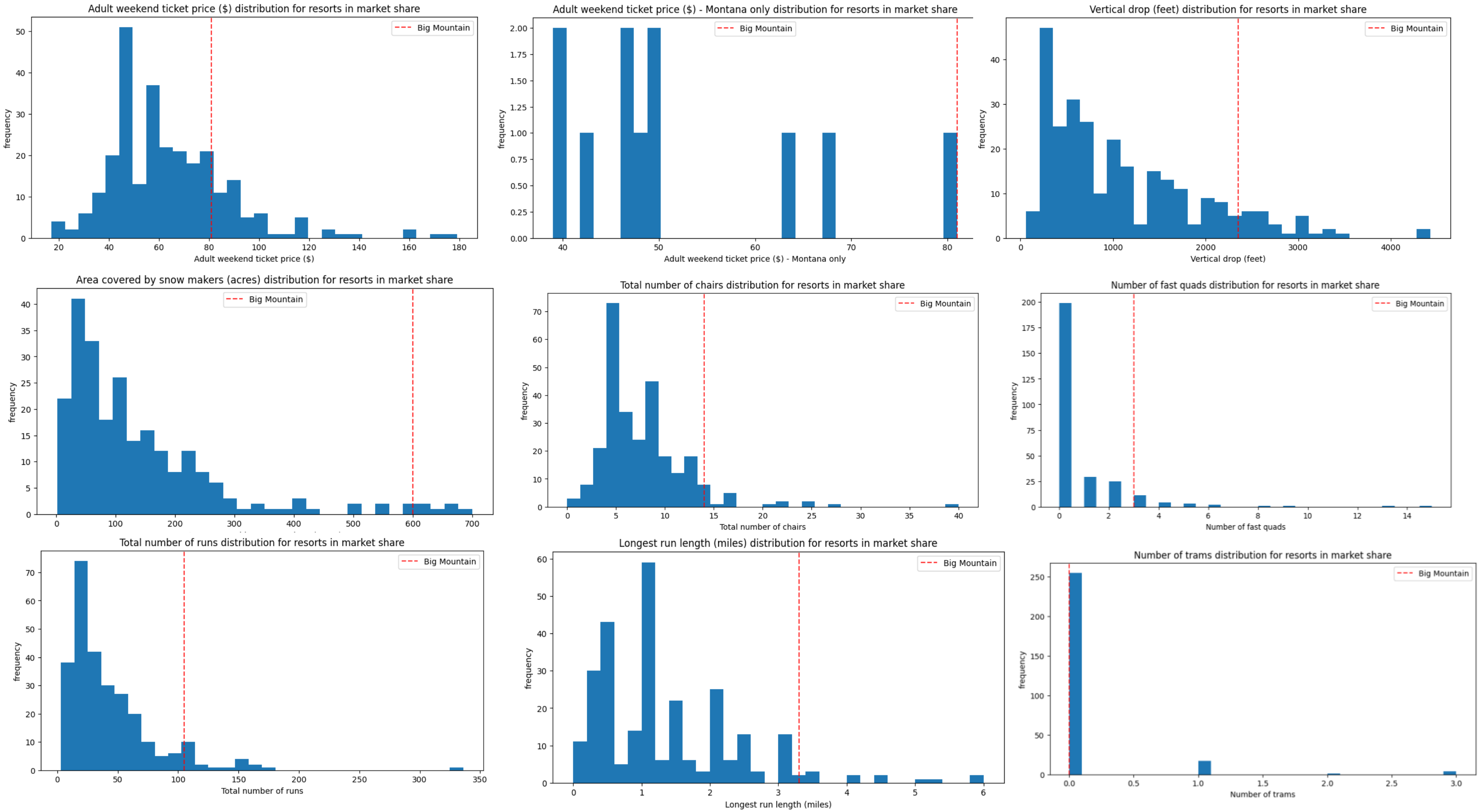
Features correlating to Price in order of importance

Feature	Corr. to Price
vertical_drop	10.767857
Snow Making_ac	6.290074
total_chairs	5.794156
fastQuads	5.745626
Runs	5.370555
LongestRun_mi	0.181814
trams	-4.142024
SkiableTerrain_ac	-5.24978

Modeling

BM’s facilities

Before modeling with a RF, the distribution of various parameters was explored to see where BM is positioned.



- BM seems competitively positioned as far as facilities are concerned, with possible improvement to vertical drop

Modeling

Random Forest Regression Predictor

- A RandomForestRegressor was applied to the training data with a Grid Search cross validation of k=5.
- Imputation strategy was ‘median’ and scaling was found to not be very impactful, so was discarded.
- Score: 0.71 with a mse= .065

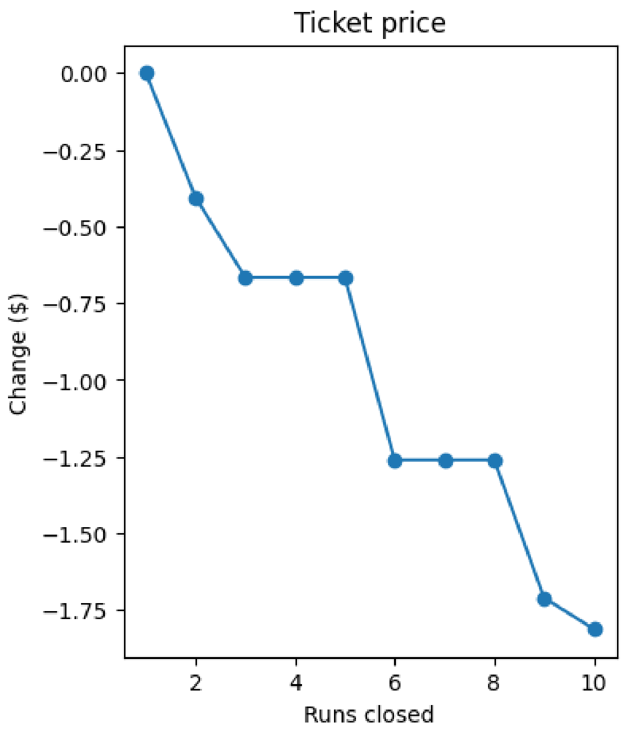
Final Model Selection (after using GridSearchCV)

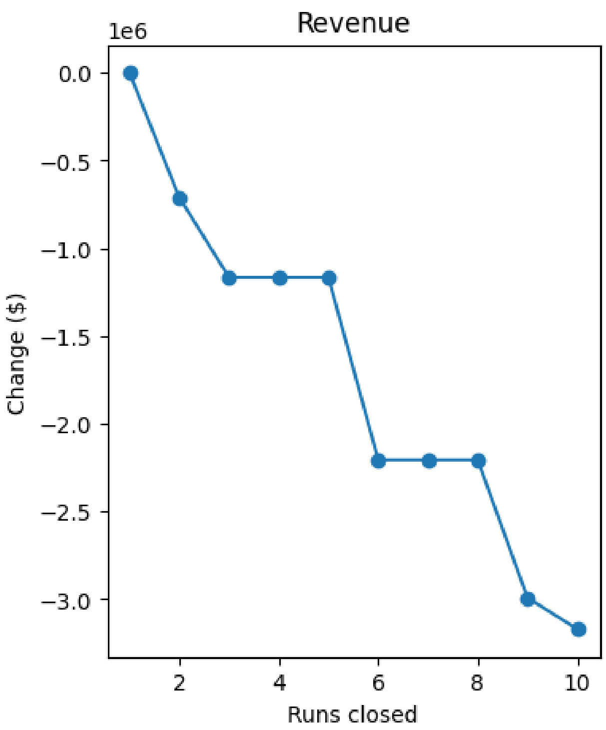
Parameter	LR	RF
MAE	11.8	9.53
MAE_mean	10.5	9.64
MAE_std	1.6	1.35

- Random Forest shows a lower mean error, as well as lesser variation in MAE. For these reasons, the RF model is selected

Modeling scenarios (350K visitors, each five times per season)

1. Permanently closing down up to 10 of the least used runs.





Will result in reduced ticket prices of $1.75, which will result in lost revenue of $325K

Modeling scenarios

2. Increasing vertical drop by 150 ft.

- This scenario increases support for ticket price by $1.99
- Over the season, this could be expected to amount to +$3474638

3. Increasing vertical drop by 150 ft. and adding 2 acres of snow-making

- Negligible change

4. Increasing the longest run by .2 miles and guaranteeing its snow coverage by adding 4 acres of snow making capability.

- Negligible change

Conclusions

- Adult Weekend ticket price is the best parameter to model around
- Random Forest Regression gives the best estimates
- Of the four avenues explored, reducing runs would result in significant loss in revenue, while increasing the ticket price would result in an almost equal gain in revenue. Based on the analysis, it seems feasible to increase the ticket value to $1.99 for a gain in revenue of $0.347M
- This still won’t be enough to cover the cost of the additional chairlift of $1.54M
- The model can be made available for serving to play with different parameters, as well as additional information such as number of visitors per season to each resort can help refine the model