Problem Statement Worksheet (Hypothesis Formation)

The Big Mountain resort is a ski resort located in Montana, 350,000 people visit every year to ski or snowboard the big mountain. They recent addition of chair lift increasing their operation cost to $1.54M per season. Need to find a way to reduce the operation cost, and project the revenue of the season based on the reduced operation cost.

H

1 Context

Big Mountain Resort has recently installed an additional chair lift to help increase the distribution of visitors across the mountain. This additional chair increases their operating costs by $1,540,000 this season. The resort's pricing strategy has been to charge a premium above the average price of resorts in its market segment. They know there are limitations to this approach. There's a suspicion that Big Mountain is not capitalizing on its facilities as much as it could.

2 Criteria for success

Reduce the operation cost without affecting previous profit margin of 9.2%.

3 Scope of solution space

The solution we are deriving the problem only for big mountain resort with addition chair lift and people visit in mind, it may not work if visitor increased dramatically.

4 Constraints within solution space

- Basing their pricing on just the market average does not provide the business with a good sense of how important some facilities are compared to others. This hampers investment strategy.
- We should increase revenue by reducing the operation cost without affecting previous profit margin of 9.2%, however it’s really hard to achieve exactly that we want.

5 Stakeholders to provide key insight

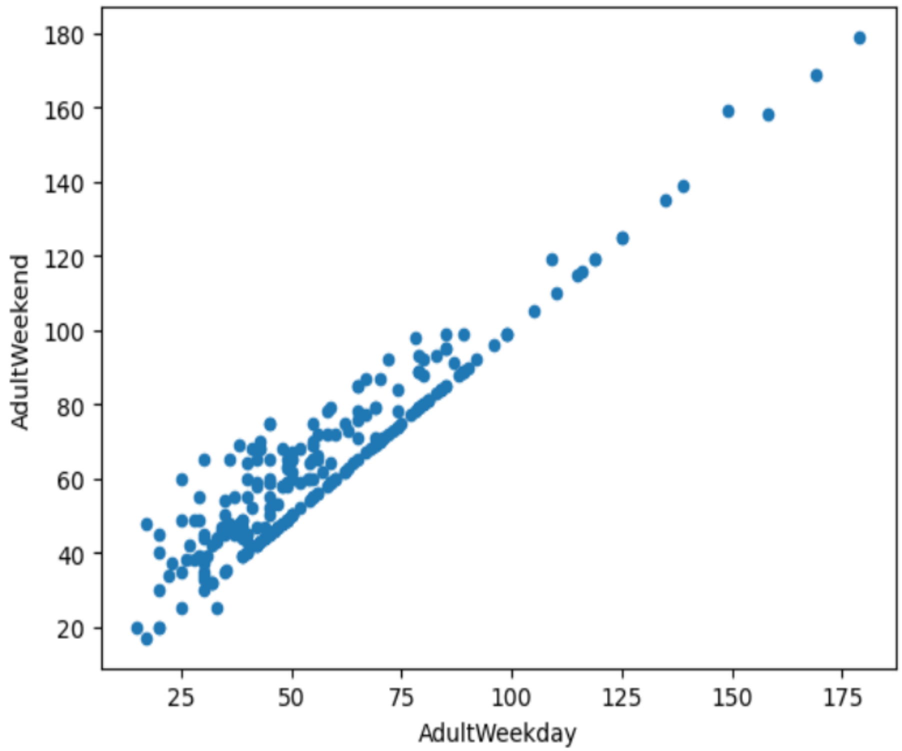
- Jimmy Blackburn, Director of Operations
- Alesha Eisen, the Database Manager

6 Key data sources

- CSV file that I got from 330 resorts in the US.

Recommendation and key findings

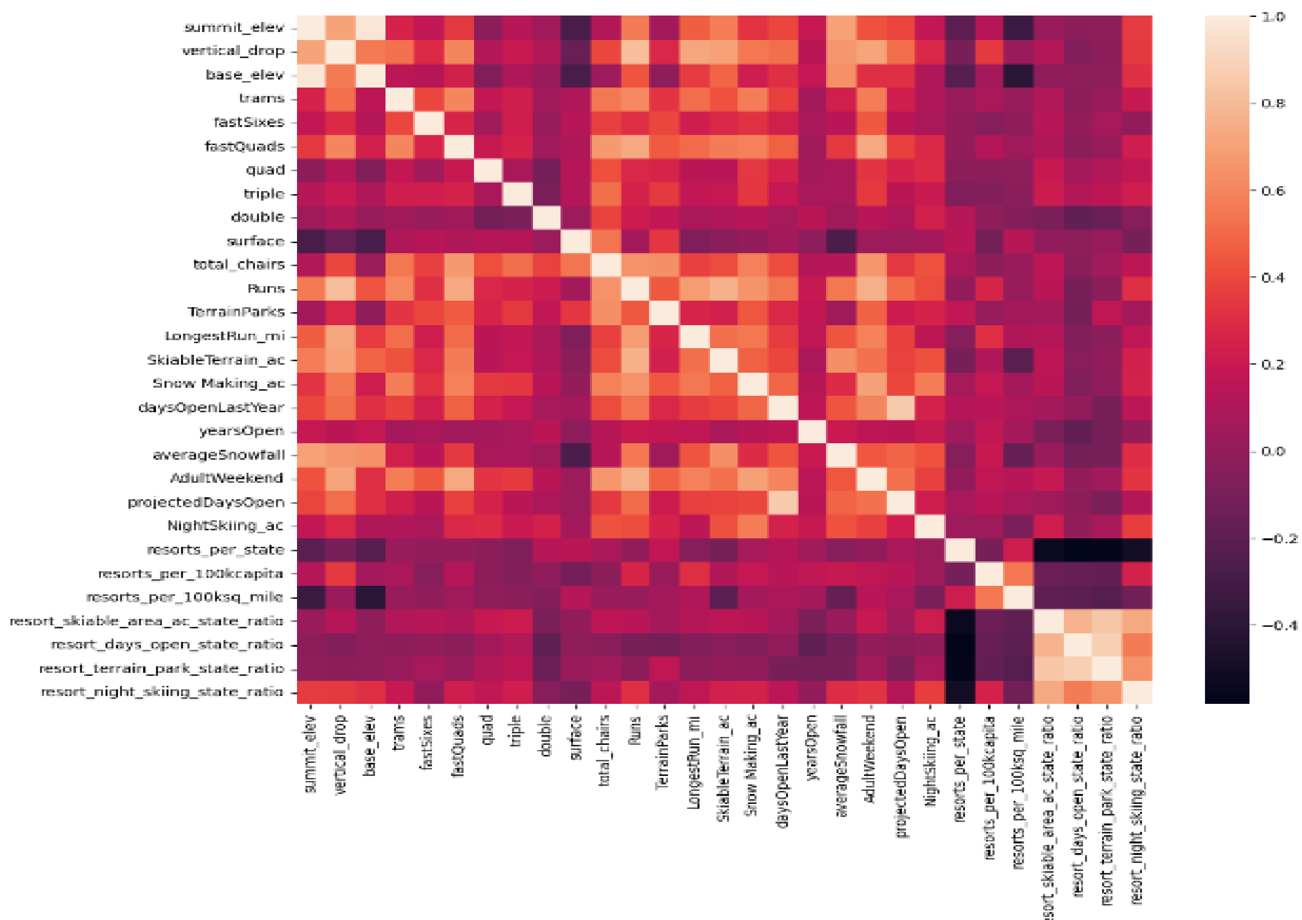
The first values inspected were the Adult Weekend vs Adult Weekday prices, was it actually advantageous to have a different price for the weekend? Most states had the same price for both, as seen by the chart below:



Adult Weekend had several missing values. Because of this, the Adult Weekend column was removed. In addition to Adult Weekend, the fastEight column was dropped because value of its values were null and the other half were mostly 0. Besides those two big ones, there were some smaller columns that had to be dropped and a lot of missing values that had to be taken care of. Once this was finished, we were left with 277 of the original 330 rows.

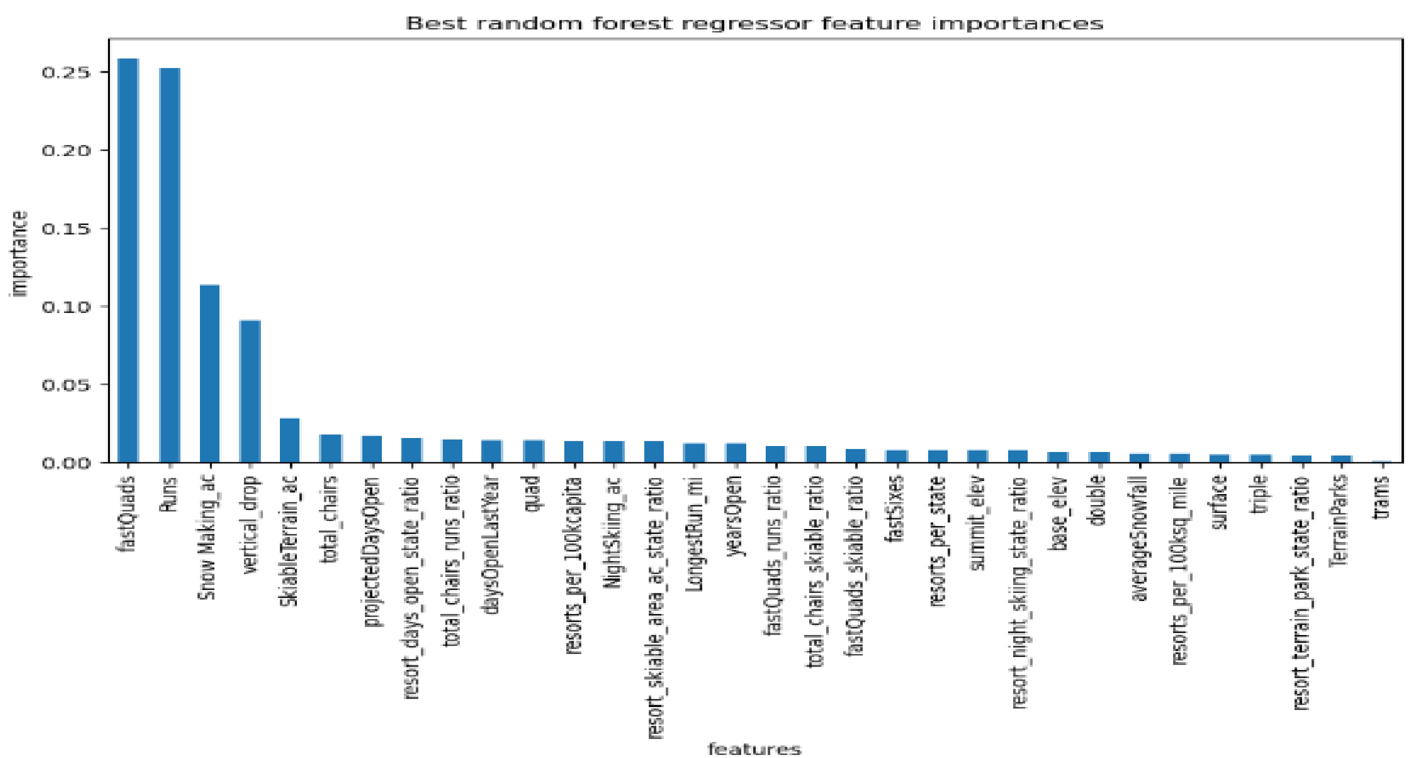
Recommendation and key findings

I wanted to look into was the relationship between the components like vertical drop, years open, or skiable areas versus the price in each state. This required a Principal Cumulative Analysis (PCA), which showed that the first two components account for 75% of the variance, and the first four account for 95%. Focusing 3 on just the two components, I scaled the data and added the average ticket price to a scatter plot. I needed to see a clearer relationship between price and components, so instead I created a heatmap to better visualize the relationship between each feature.



Recommendation and key findings

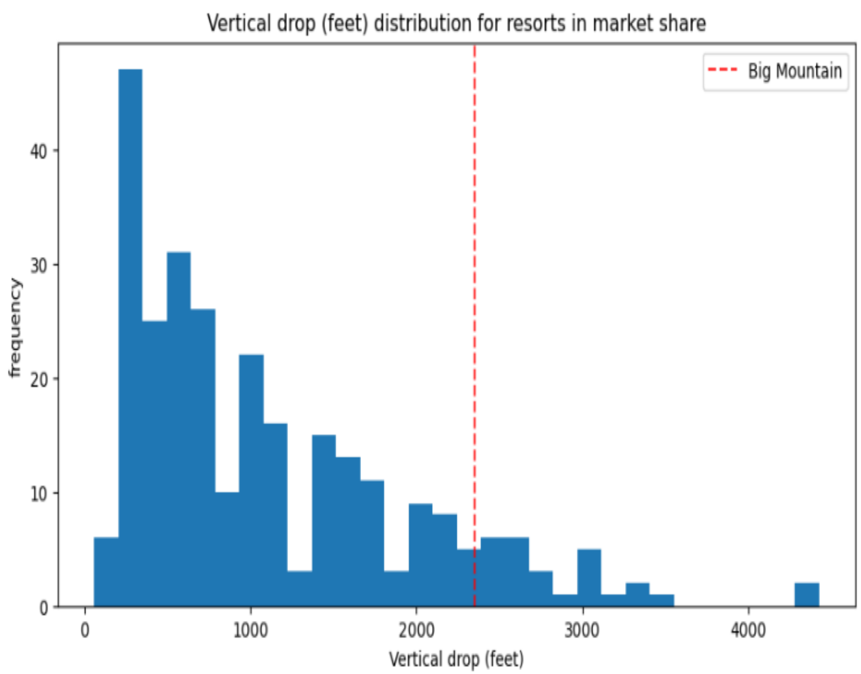
After identifying the four categories with the strongest correlation to price, the first step was to take an initial average as a “best guess” for pricing. This resulted in $83.81 initially which was the baseline I used to compare other prices to moving forward. Mean wasn’t the best for this because our Absolute Error was off by around $19, much too large for a price like this. Instead, I performed a regression using the median between results. The Mean Absolute Error was only off by $9 this time, but I knew this could still be improved. I also decided to create a data pipeline in order to efficiently produce identical results to make comparisons easier. The next regression was based on a Random Forest Model which helped identify that imputing the median value helps with the MAE of our four components. In addition to our four components, during the analysis I found that vertical drop also plays a big role in determining ticket price. Once this new component was added to the Random Forest Model, the Mean Absolute error was down to around $1 which was an acceptable amount of variability.



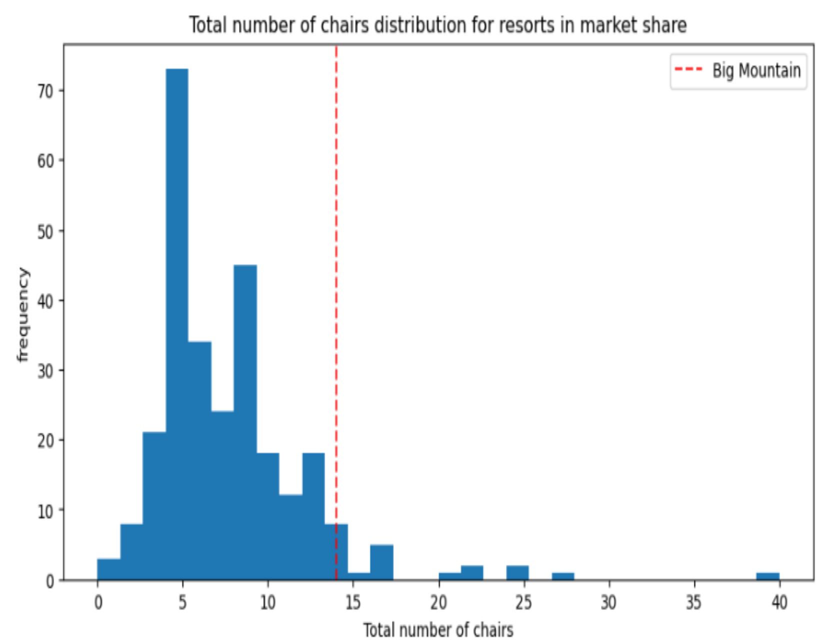
Modeling results and analysis

At this point I have picked the top components and a method of regression. I can use these together to create a model that actually gives us a data-based ticket price. To make the model as accurate as possible I boosted the number of components to eight by adding total_chairs, LongestRun_mi, trams, and vertical_drop to the list. To determine a fair price, I needed to see where Big Mountain Resort (represented by the dashed red line) ranked in these categories. - Vertical Drop - Total Chairs 6 - Fast Quads - Longest Run - Number of Trams - Number of Runs 7 - Area Covered by Snow Makers - Skiable Area.

Vertical Drop

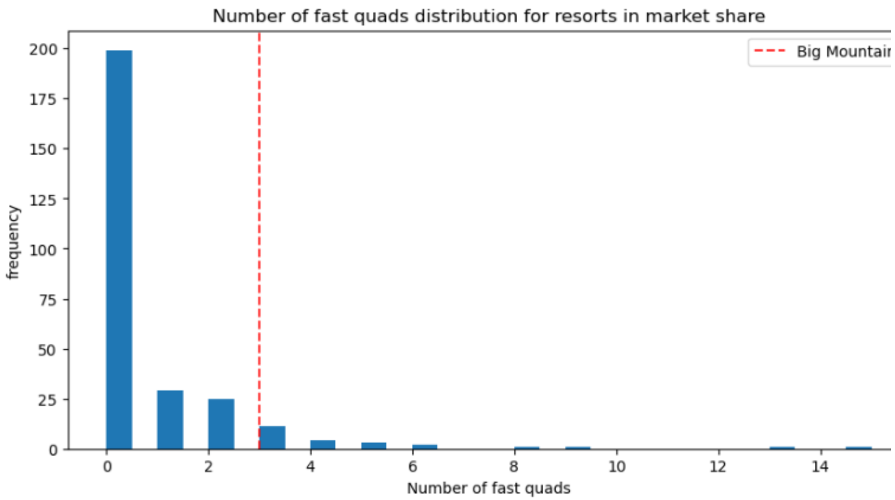


Total Chairs

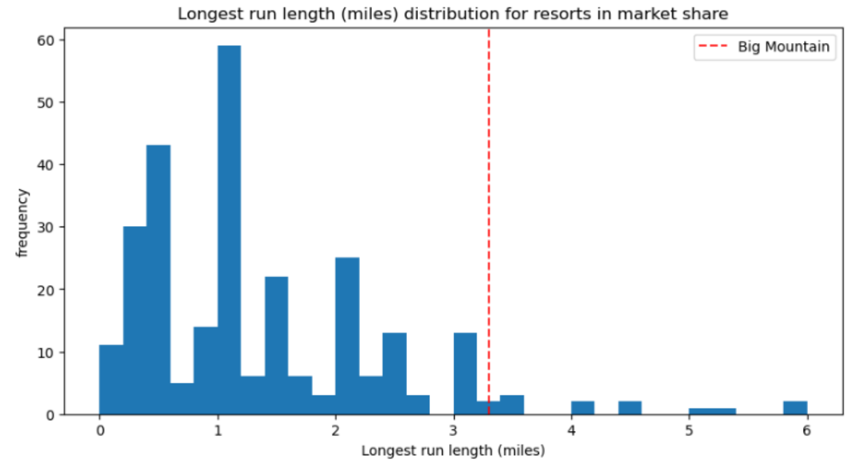


Modeling results and analysis

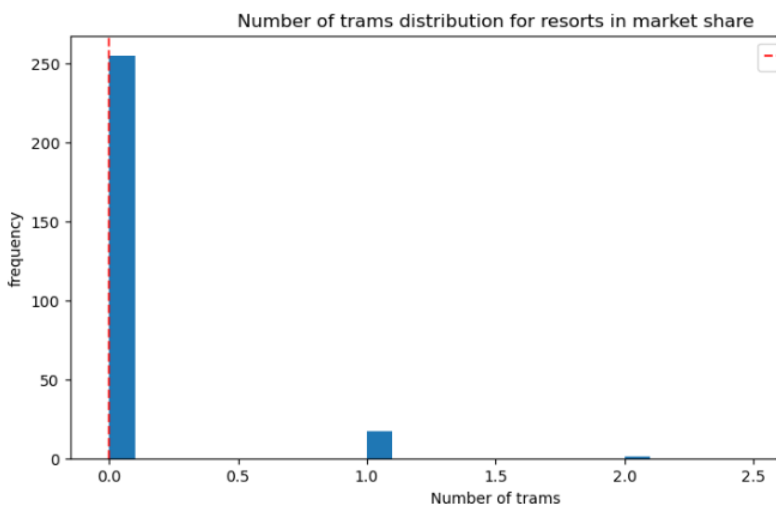
Fast Quads



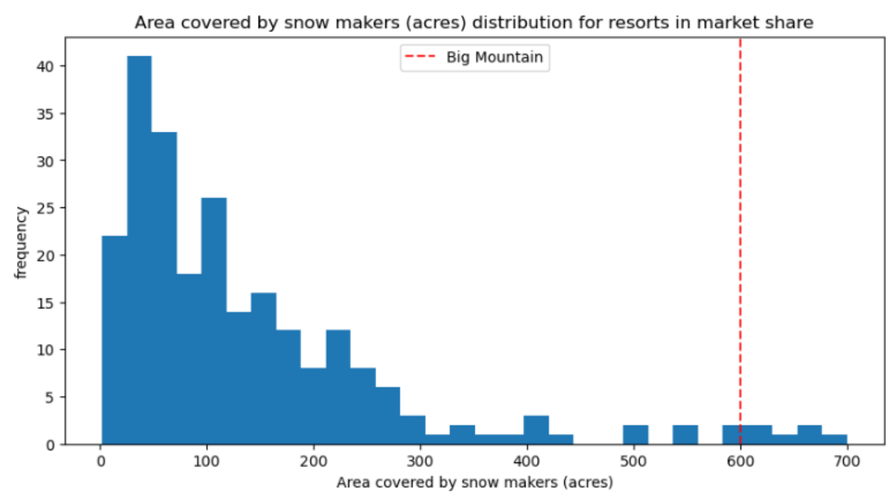
Longest Run



Number of Trams

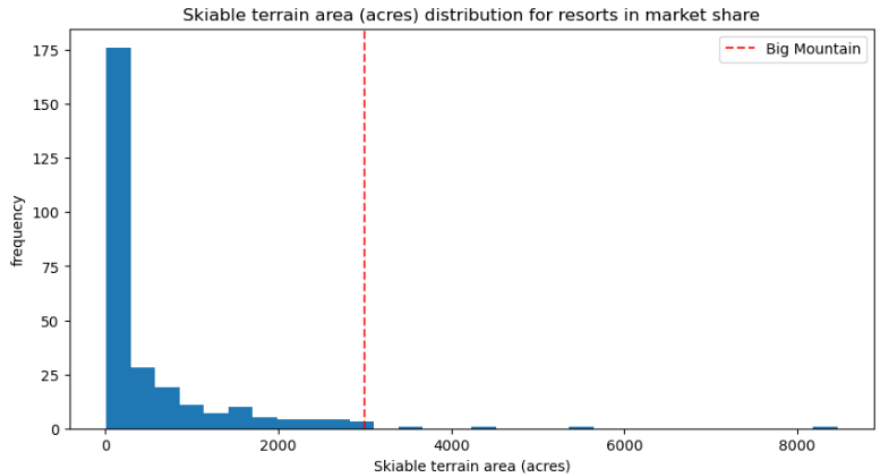


Area Covered by Snow Makers



Modeling results and analysis

Skiable Area



As you can see, Big Mountain Resort either ranks high or is well above the average in each category. With the exception of trams which most resorts don’t have anyway. All this shows us that Big Mountain Resort is an exceptional resort with many great facilities, and the price should reflect that. With that we get a modelled price of $96.62, which is well above their current price of $81.00.

Summary and conclusion

The Random Forest Model placed Big Mountain Resorts ideal ticket price at $94.22, much higher than their current price of $81. Therefore in order to increase revenue and lower operation costs, Big Mountain Resort should:

•Raise ticket price to $89.99

•Have 4-6 Runs closed each day to save on operations cost

To improve their facilities, encourage higher ticket prices, and improve their resort's overall appeal:

•Increase vertical drop (a highly price correlated feature) by 150 ft

•Add one chair lift

•Add an additional run

•Add 2 acres of snow cover