# Big Mountain Resort Report

**Introduction:**

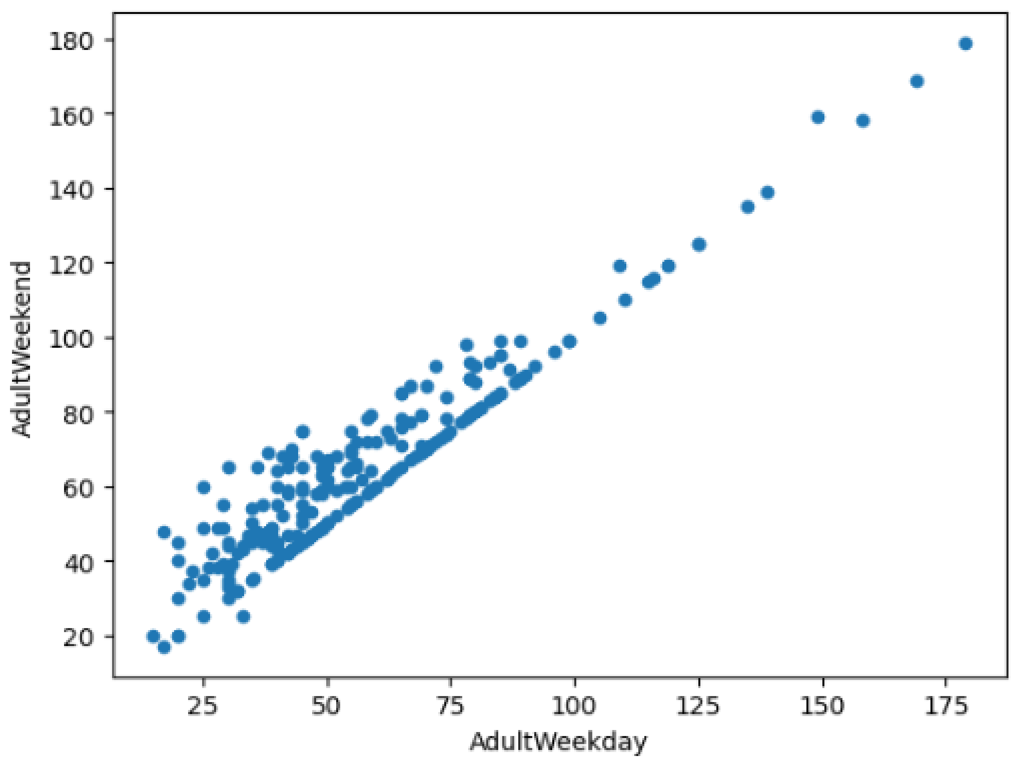
Big Mountain Resort, a ski resort located in Montana, has recently installed another chair lift to aid in increasing the distribution of its visitors across the mountain. This installation has increased their operating cost by $1,540,000. This increased expense has caused Big Mountain Resort's leadership to pursue a different business strategy to accommodate the increase in their operating cost.

**Problem Statement:**

How can Big Mountain Resort adopt a data-driven pricing strategy to recoup the increased operating cost of 1.54MM and maintain or increase their profit margin for the season over the next year?

**Data Wrangling:**

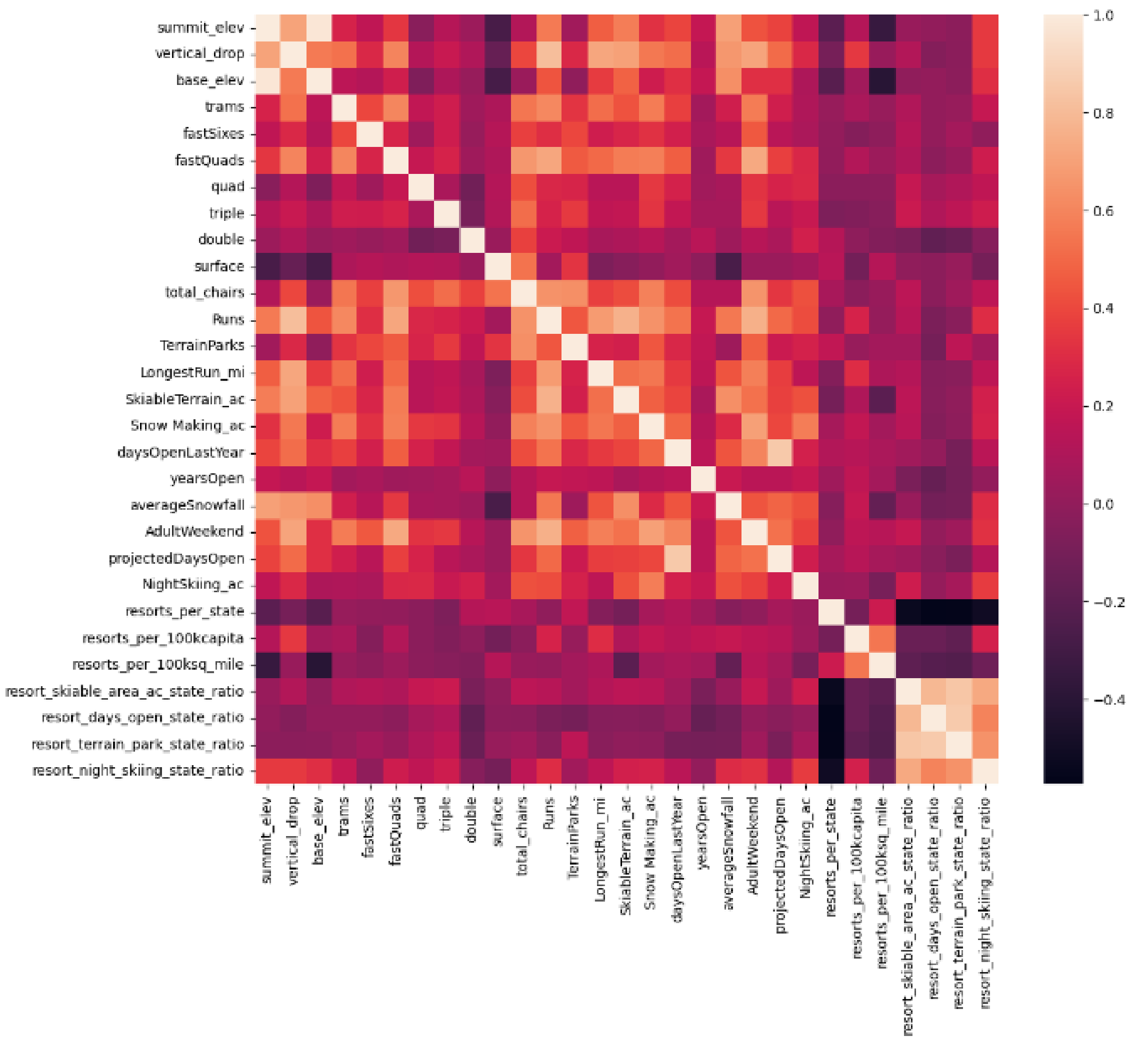
The first values inspected were the AdultWeekend vs AdultWeekday prices. The chart below shows that most state had the same price for both the weekend and the weekday:




With combination of both prices being the same and the AdultWeekend column missing serval values, the AdultWeekend column was removed. We also removed the fastEight column because a lot of the values were null while the other values were stated as 0. A few other rows were removed because of inaccurate information. Once this process was finished, we had 277 rows left out of the original 330 rows.

**Exploratory Data Analysis:**

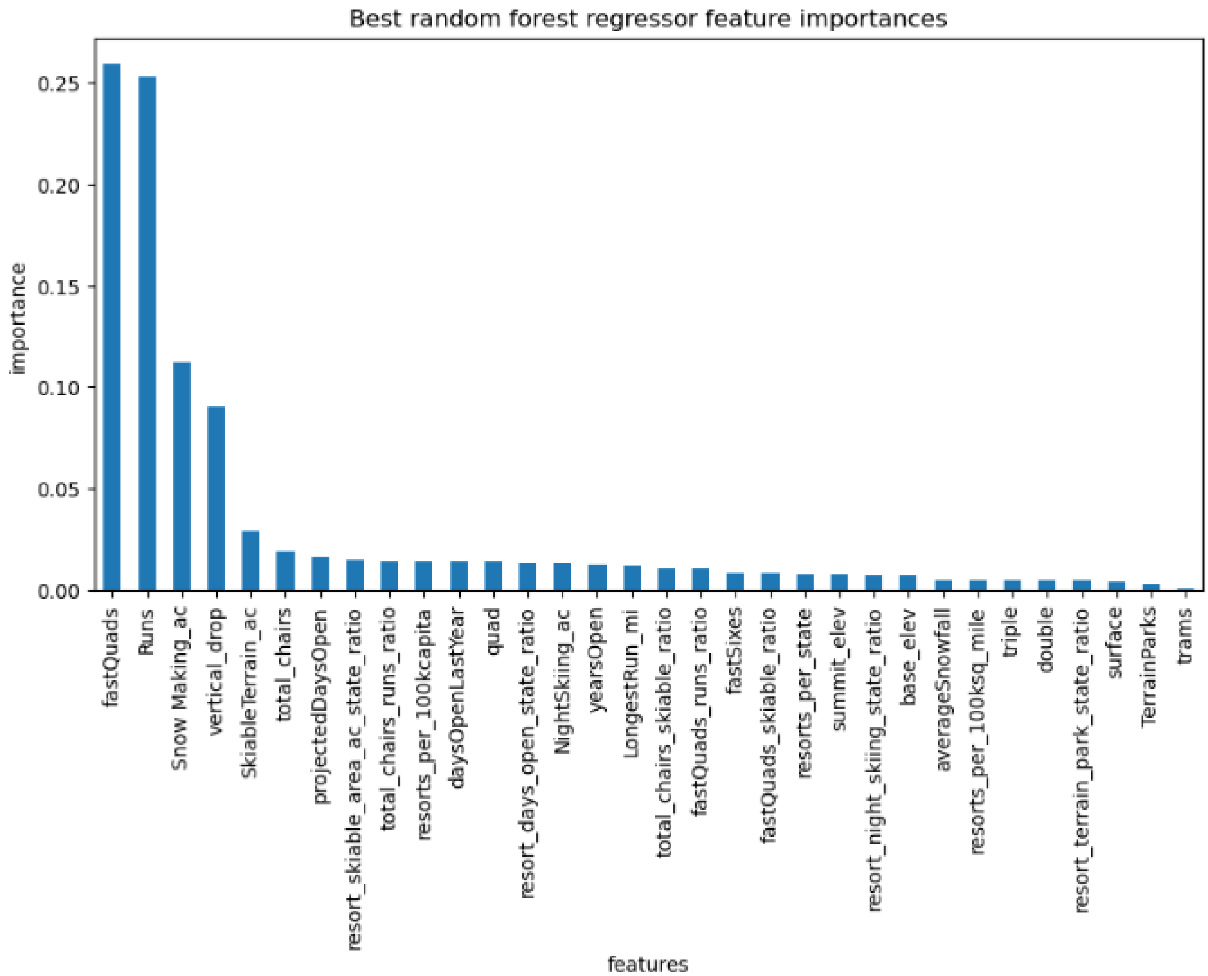
In this stage we explored the data to find actionable insights and patterns.  From this analysis I found that the largest correlation can been seen between the ticket price and the features in each resort. The heatmap below was created to better visualize the relationship between each feature:




For example, there is a positive correlation between ticket prices and features like fatQuads, runs and total chairs. With these correlations and data, we can estimate the ticket price for Big Mountain Resort based on the features present in the resort.

**Pre-Processing and Training Data:**

We initially started off with a best guess price of 63.81 and used this figure to compare with the other outcomes for pricing. Once the Mean Absolute Error was analyzed, the average of known values model would be off by 19 dollars each time. This was too large for our price point. We then used a linear regression model. In this process, we assigned median values found in each row to the missing features. We then ran the tests and noticed a big improvement. The Mean Absolute Error test reduced to only around 9 dollars off the known values. In the next step, we created a data pipeline to effectively produce similar results to make comparisons easier. We also used a Random Forest Regressor. From running these tests, we noticed that the biggest impact on price on average was the vertical drop. Once this new component was added to the Random Forest Model the Mean Absolute Error was down to an acceptable amount of about 1 dollar.

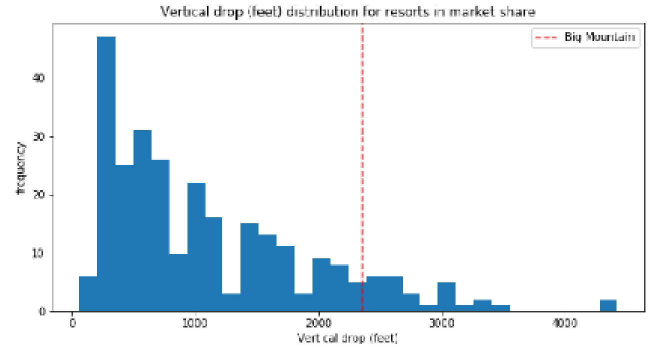



**Modeling:**

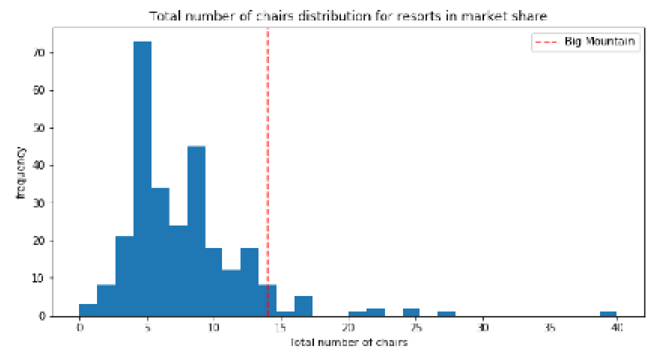
In this stage we picked the top components and the method of regression. We must use these together to create an accurate model that gives us a data-based ticket price. To accomplish this, we had a list of 8 components (Vertical drop, Total chairs, Fast Quads, Longest Run, Number of Trams, Area Covered by Snow Makers and Skiable Area). We then tried to see where Big Mountain resort ranked in these categories.

Big Mountain resort ranked high or above average in each category except for trams. Some of the components can be seen in the diagrams below:

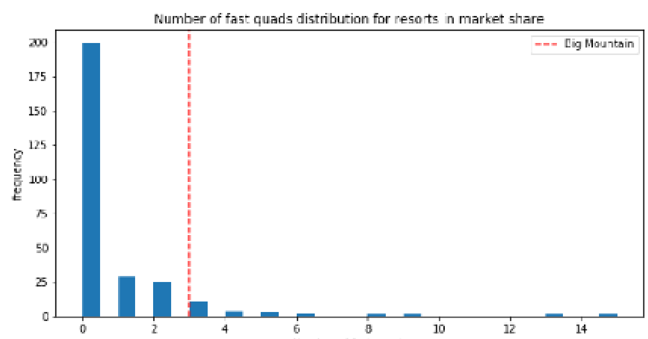
- **Vertical Drop**



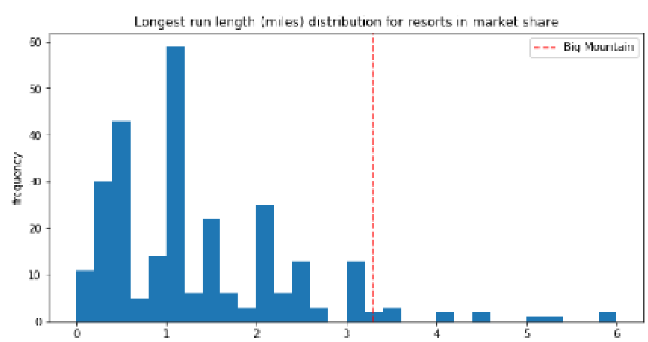

- **Total Chairs**




- **Fast Quads**




- **Longest Run**

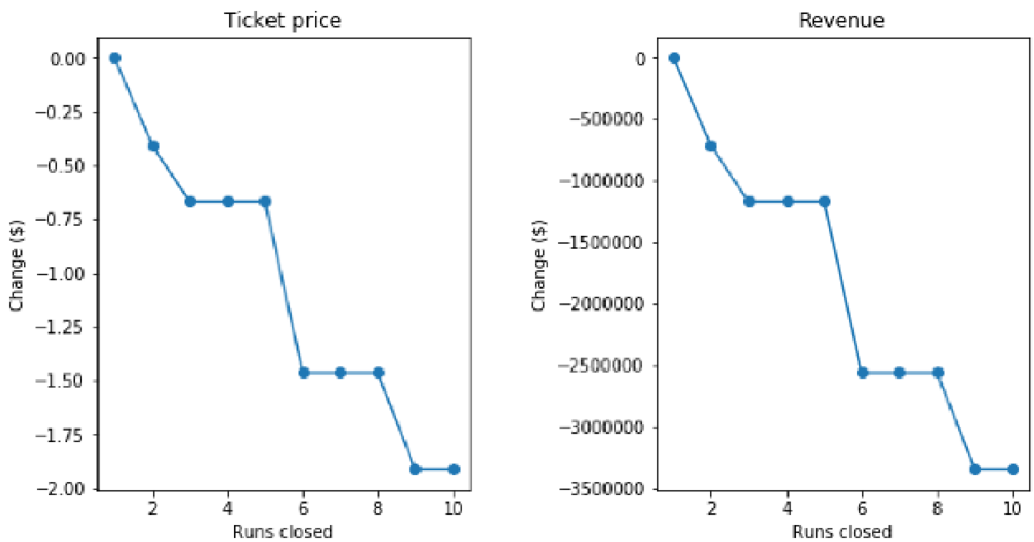


This analysis further justifies a price hike as Big Mountain resort is a premium resort with exceptional features. Taking all this into account, we get the modelled ticket price of $97.29.

**Conclusion:**

Big Mountain resort is currently undervaluing their ticket price by charging $81. Based on the features and services that are available in the resort, the model suggests that they should increase their ticket price to somewhere around $97.29.

In addition, Big Mountain resort can also reduce their operating cost by closing 4-5 runs that are seldom used. The model indicates that keeping 4-5 runs closed only has a small effect on revenue.




There were some deficiencies in the data that limited this work. We are limited to only the tickets prices of the other resorts. We didn't really have information for other factors that may affect the ticket prices.

This model should work well for their current goal of increasing profitability. It should also be useful in the future. If additional information is provided, it should be easy to include and create new features which will only make the model more efficient.