Big Mountain Resort Project

Step 1 - Problem Identification:

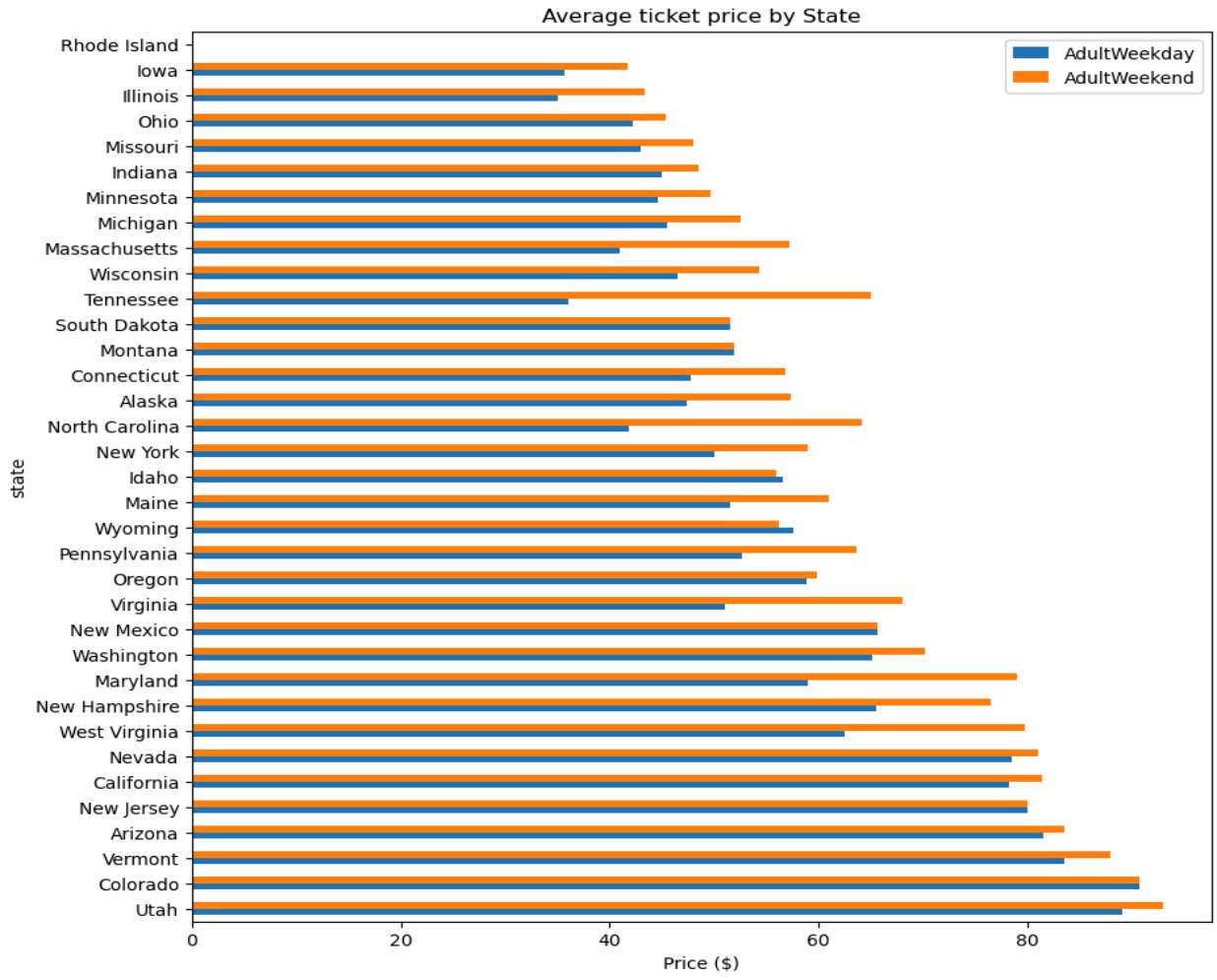
In this project, I was able to practice my coding skills. I completed the Data Science Method steps from beginning to end. In the first step, I was given information on the increase off $1,540,00 for operation cost due to an addition of a chair. The management team is looking into cutting costs without undermining the ticket price or finding support for an increase in ticket price. I chose the latter. With all the amenities that Big Mountain Resort has to offer, the team should capitalize on it.

Step 2 - Data Wrangling:

In the second step, I collected, organized and cleaned datasets. I loaded the ski data csv, and used the .info() and. head() method to retrieve a quick summary of the data. I explored Big Mountain Resort's data, and there was no missing data, but many missing values in other resorts. 'FastEight' had around 50% missing, and 'AdultWeekday' and 'AdultWeekend' had 15-16% with the latter having a few more records. I also explored the 'Region' and 'State' columns, and plotted the data. Montana came in 13th for most resorts while New York came in first. Then, I used a bar and boxplot to compare ticket prices by state, and found Montana's ticket prices to be in the middle. Also, there is low variability for Montana's weekday and weekend ticket prices, both are $81. Next, these feature values, 'TerrainPark', 'SkiableTerrain_ac', 'daysOpenLastYear', and 'NightSkiing_ac' , are some that stood out to be explored further. I dropped the 'fastEight' column and rows missing price data. Since there are less missing values in the 'AdultWeekend' column, I can focus on that price, and explore ways to support a premium ticket price.

Big Mountain Resort Data:

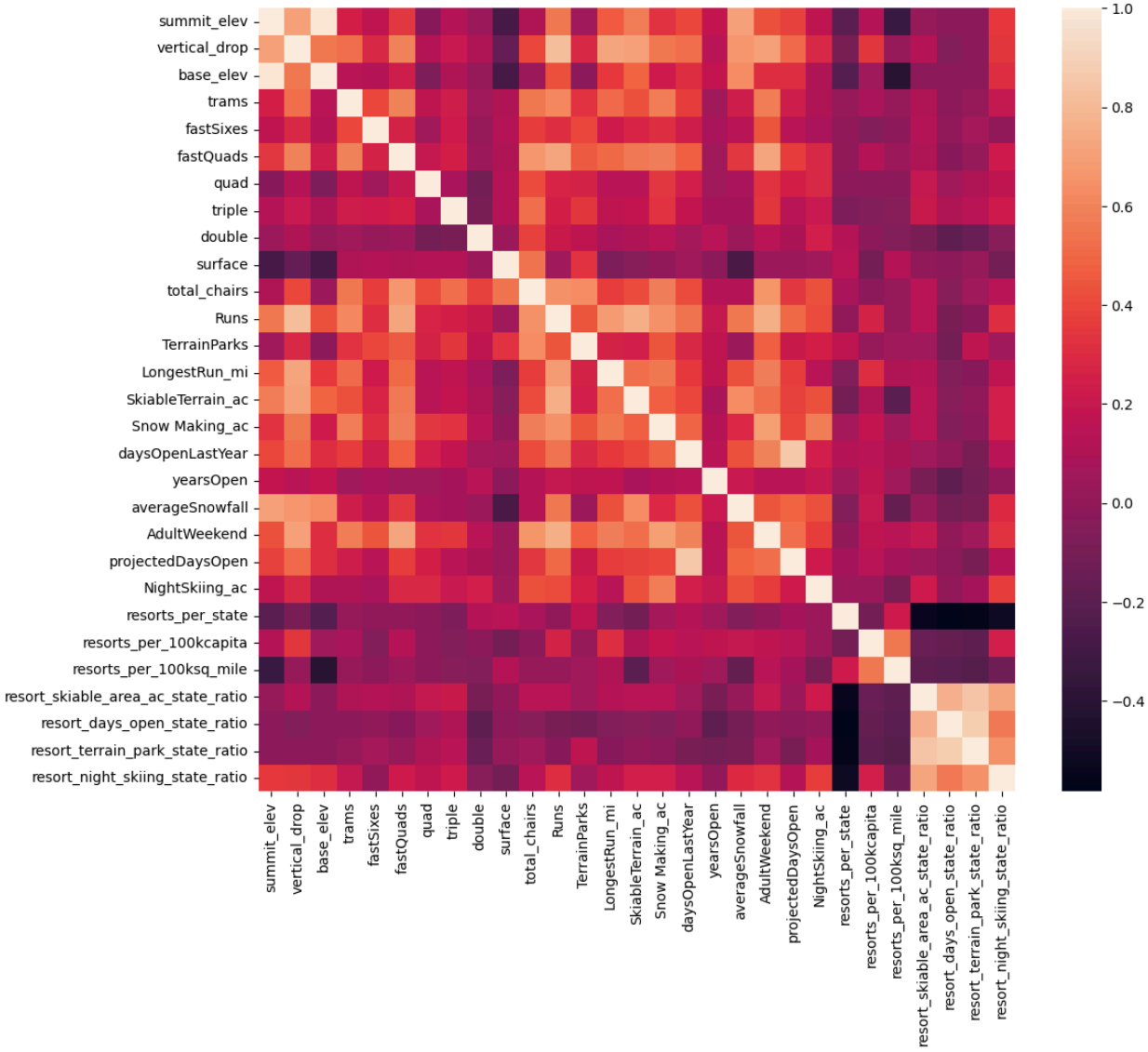
Name	Big Mountain Resort
State	Montana
Terrain Park	4.0
Skiable Terrain	3000.0
Days Open Last Year	123.0
Night Skiing	600.0
Vertical Drop	2353



Step 3 - Exploratory Data Analysis:

In the third step, I explored the relationship between state-wide data and feature values. When using the state data, it shows that Montana has the third largest state area and fourth largest for skiing, but did not land in the top five for population, night skiing, total days open or amount of resorts. To get a better understanding of the correlation of ticket prices and state data, I calculated the ratios of 100k per capita and square mile. Then, I used principle components analysis(PCA) technique to compare the original feature with the derived features and merged tables together. When visualizing the data, I used a heat correlation map and found a correlation between ticket price and five features, which are fastQuads, Runs, Snow Making_ac, total_chairs and vertical_drop.

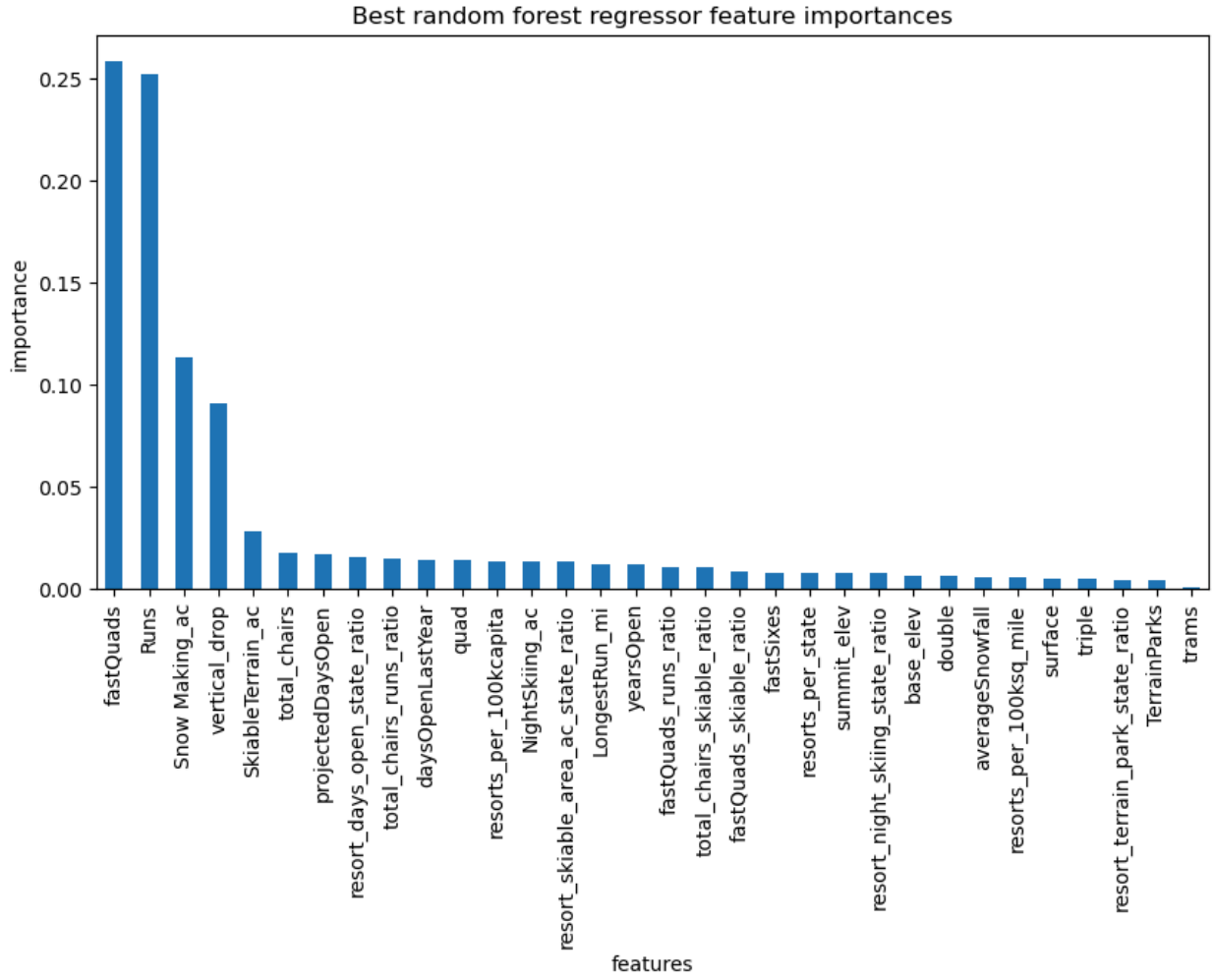
Correlation Heatmap In Ski data



Step 4 - Pre-Processing and Training:

In step 4, I modeled and trained my data by partitioning it into training and testing splits and using a DummyRegressor. I dropped the ticket price column and tested the mean and median using r-squared, mean absolute error and mean square error. The mean and median were predicted accurately. I replaced missing values by imputing the median, scaled the data and trained a linear regression model. Next, I used a pipeline from sklearn to complete all the same steps in a quicker, more efficient way. I used the function SelectKBest and cross validated

for estimating model performance. I found that the features affecting ticket price are vertical_drop, Snow Making_ac, total_chairs, and FastQuads positively, while SkiableTerrain_ac had a negative effect. Next, I used the Random Forest model and cross validation technique. The four dominant features were fastQuads, Runs, Snow Making_ac and vertical_drop. I prefer the Random Forest model because it has less variability when testing the mean absolute error.



Step 5 - Modeling:

In the last step, we took our model for ticket price and looked at what Big Mountain's facilities might support. There is support of an increase in ticket price to $95.87 with an expected mean absolute error of $10.39. The original price is $81. One reason for the underpriced ticket could be because Big Mountain's price is higher than the average price in Montana. However, with all the amenities Big Mountain has to offer, the management team should capitalize on a ticket increase. Also, we completed four modeling scenarios for either cutting costs or increasing revenue. The one that would be beneficial is adding a run, maybe for Hellfire, increasing the vertical drop by 150 feet, and installing an additional chair lift.

Conclusion:

There are 350,000 visitors per year, and on average each visitor purchases five days of skiing. The ticket price increase would cover the operating costs of $1,540,000 as well as the modeling scenario in the above paragraph. Also, Big Mountain's business analyst should learn the model so they can test new parameters. We should be going back to the model to keep up with the resort’s value and profit.