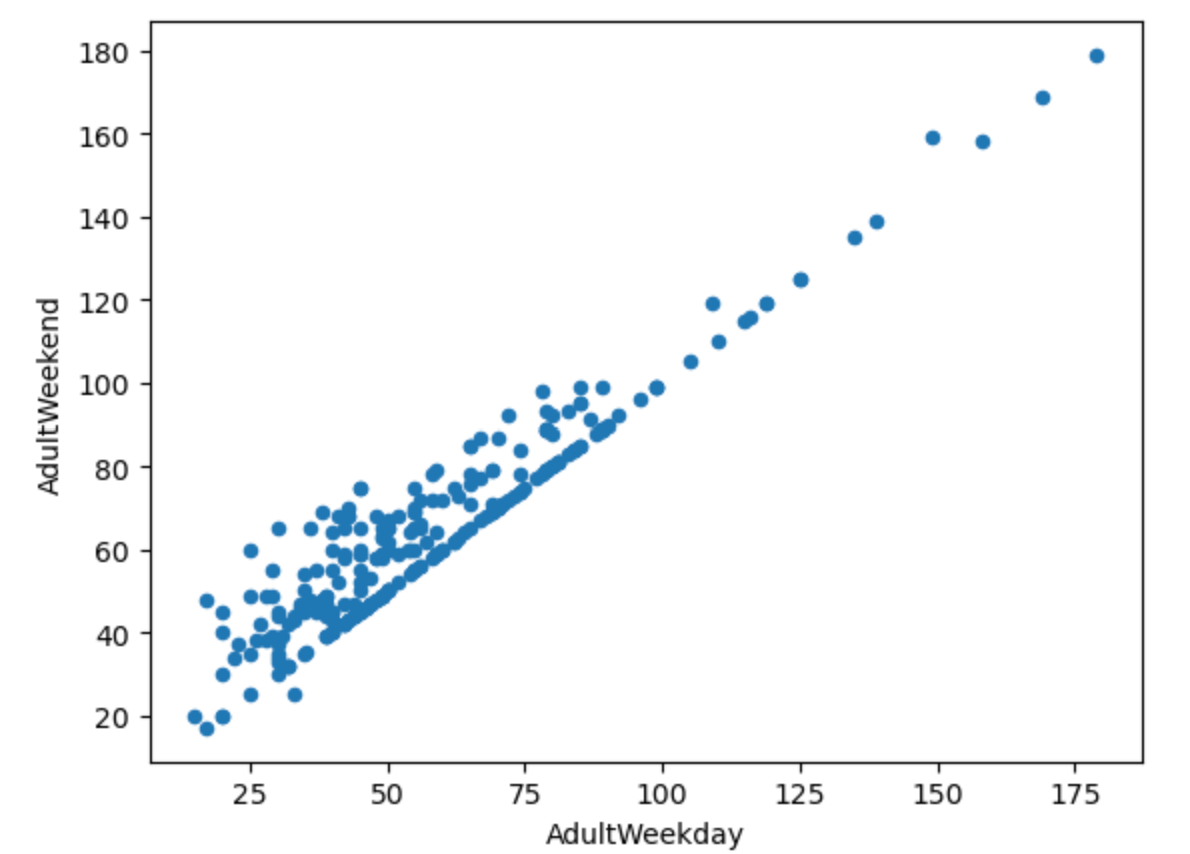
Big Mountain Resort is a ski resort located in Montana, offering spectacular views of both Glacier National Park and Flathead National Forest. Big mountain has recently added an additional ski-lift to help with the distribution of its visitors. By the end of the next fiscal year, Big mountain is looking for different pricing strategies to help combat the $1.54 million dollars of additional operating costs.

How can we achieve this? We will attempt to implement a new pricing strategy by either cutting additional costs, or by updating our facilities to best enhance customer experience. Our focus is going to be on our facilities. Through data analysis we will see where we compare with other companies in the same industry. Unfortunately, our data is a bit limited; we are restricted to data solely on terrain and lift equipment. Nevertheless, I am confident we can still achieve the desired results.

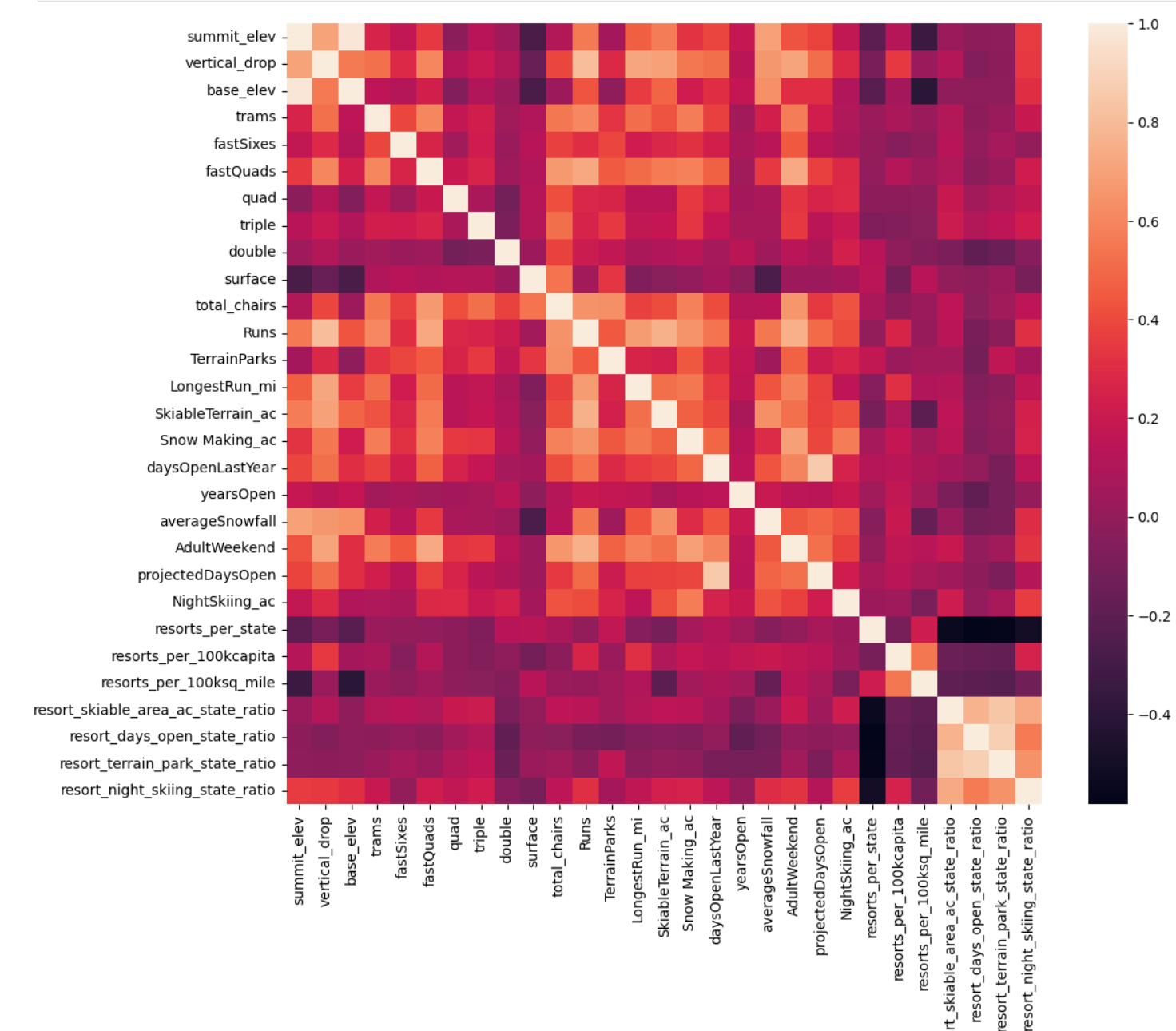
Firstly, we begin with data wrangling. We want to focus only on information that we need.

Our original data set started with 227 rows and 27 columns. We ended up removing the fastEight column entirely. It turned out over half of the data on that column had null values, and the other half were missing entirely. Because we are focused here on ticket pricing, we ended up removing all data where none were available to us. Our data set gave us access to both weekday and weekend pricing. We created a scatterplot to analyze the relationship between the two.



It turns out the relationship is fairly linear. Because the weekday prices had more missing values, we decided to remove that column as well. At this stage, we still haven’t decided what to do with our state data, whether to focus only on resorts in Montana; more is to be determined.

In the next part, we will explore our data some more. We attempted to see the correlation between state and ticket price, however we were unable to find any meaningful relationship between the two. We also wanted to see which features most influenced the ticket price. We did this by using Seaborn’s heatmap function.

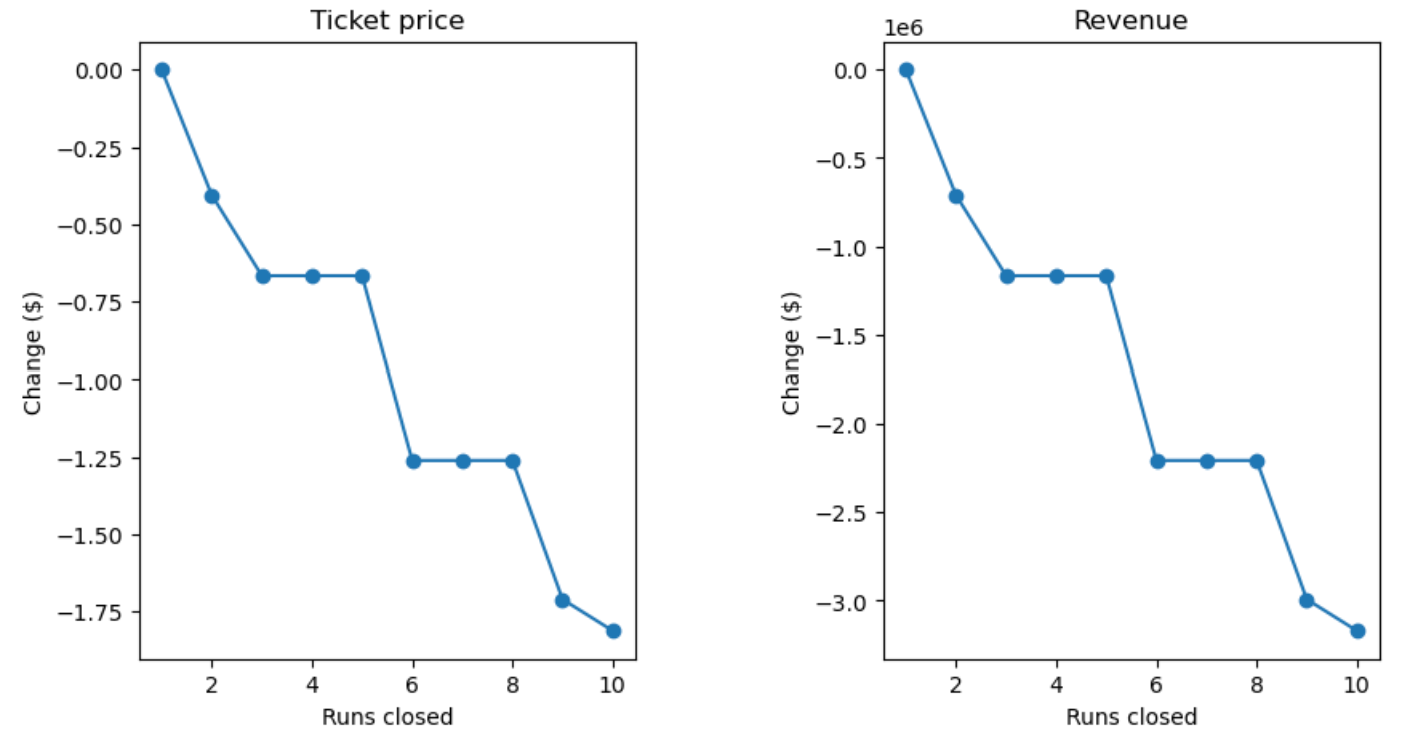


The image above not only shows our potential target features, but also the relationships between each feature in our entire data set. With regards to ticket price, a few stood out including: fastquads, total number of runs and even snow making accumulation. This information will be very useful in future modeling.

Now it is time to focus on the preprocessing stage. We began by separating our data into 2 parts, a training set and a test set. We created a few models as well. Our first model used only the average price as a predictor, the second was a linear model, and the third was a model using a random forest regressor. Calculating for Mean Absolute Error, testing all 3 models, the random forest seems to be the winner. It was on average $1 less than the linear model, and showed less variability in the results.

Finally, we can get to the modeling stage. At the moment, Big Mountain charges $ 81.00 for their ticket prices. Our model suggests we may be underpricing. By adding an additional lift and increasing vertical drop by 150 feet, we can add support for increasing prices by $8.61. We also looked into closing down the 10 least runs used.

Looking at the image below:



We can see that closing down a single run had no effect on the price, but anything more began to lose support for an increase.

Given the analysis above, in order to maximize profits for Big Mountain, they should close down their least used run, in addition to adding the new run for the next ski season.