**Big Mountain Resort Case Study**

**Nolan Long**

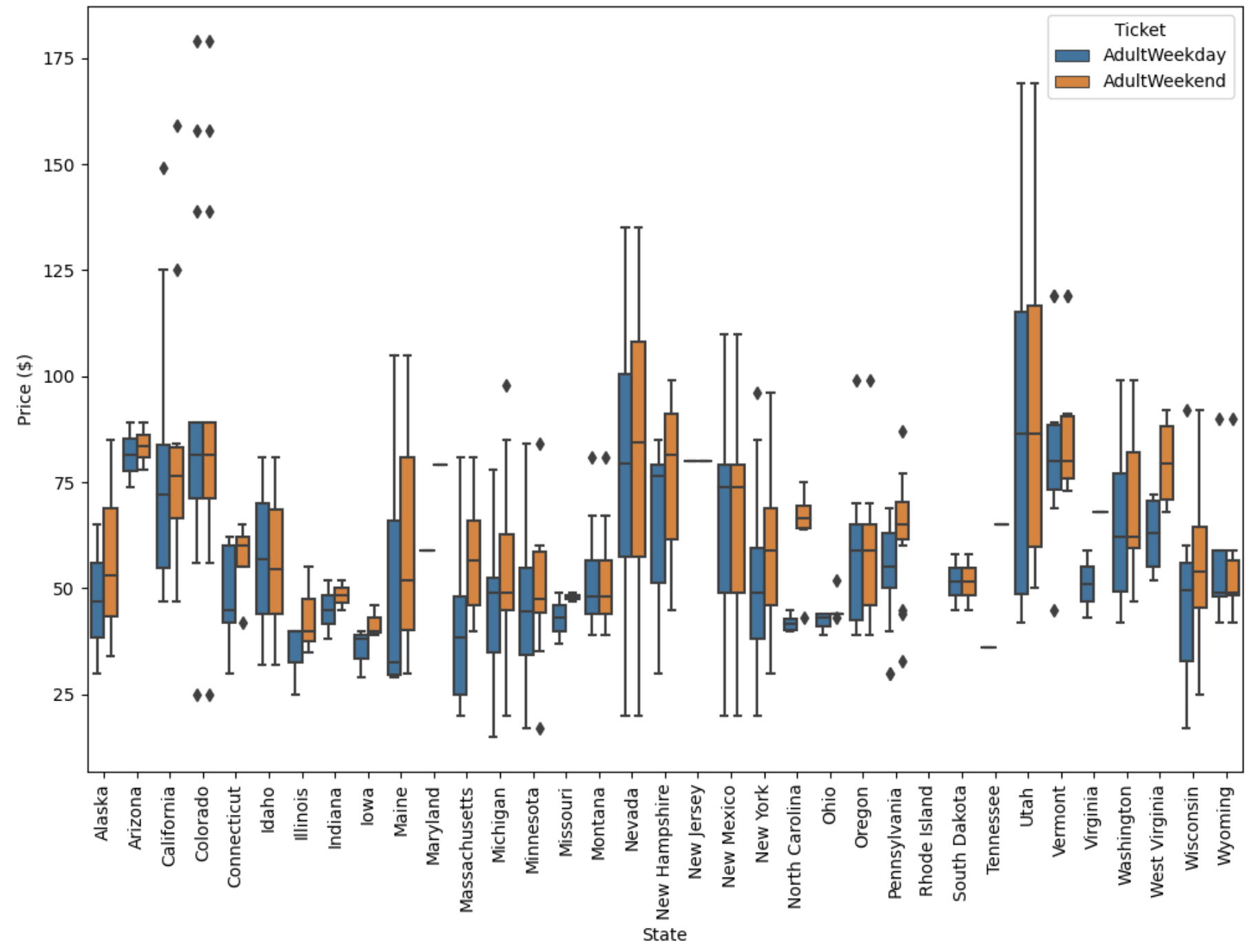
Problem Statement:

In this case study I looked at Big Mountain Ski Resort. Based in Montana this resort serves 350,000 customers a year and has 2.352 foot elevation change. I was asked to find optimal pricing for the resort as well as any changes in amenities that could increase profits.

The Data:

I was given a csv of 320 resorts along with features about the mountain such as the amount of chairlifts, total run length, and skiable acreage, and cost of a ticket. After cleaning the data by removing resorts without ticket prices, fixing same named resorts, and removing features with only a few entries I was left with complete data on 277 resorts to model from.

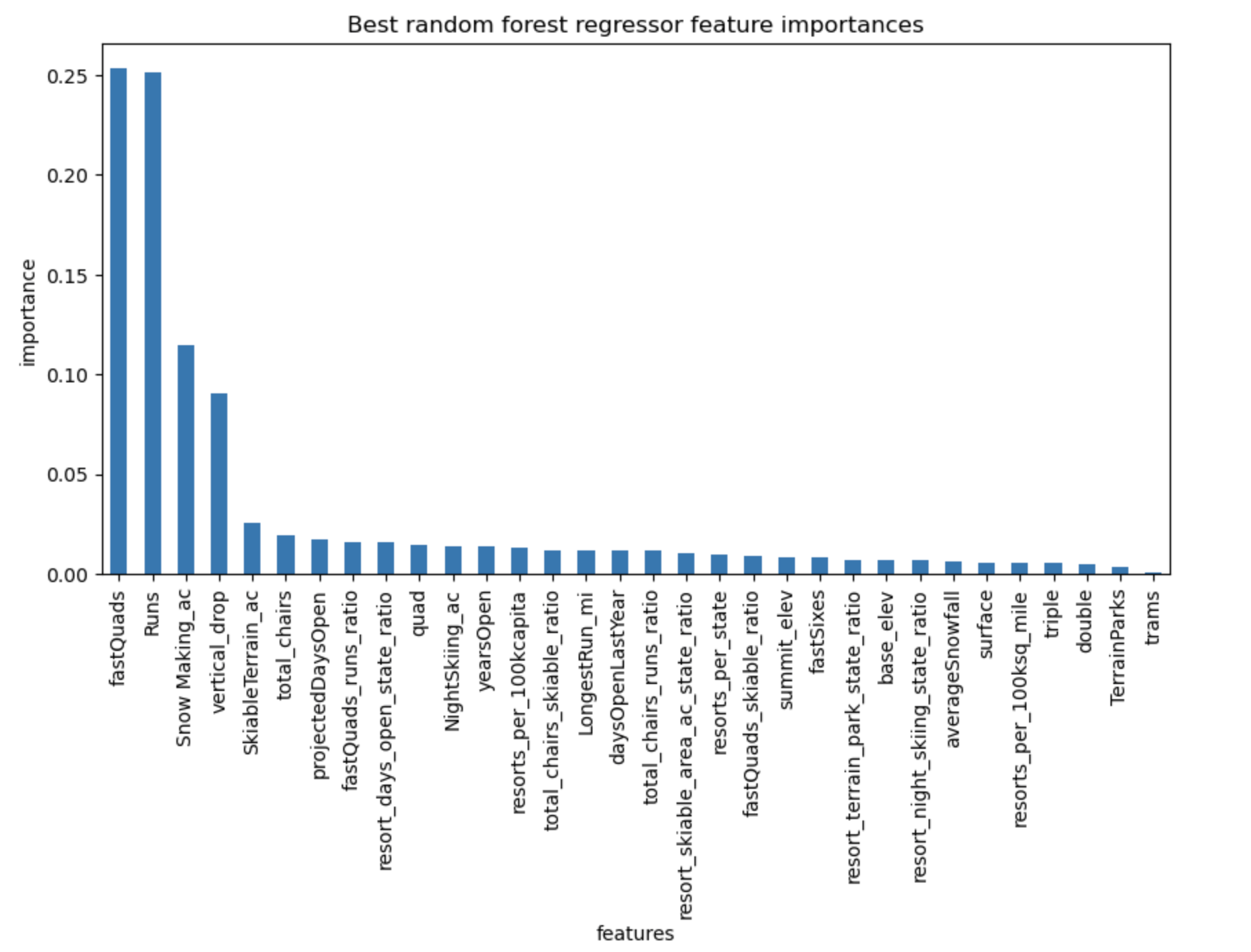
Exploratory Data Analysis and Feature Engineering:

When exploring the data I looked at a variety of features.The main categorical data was the state the resort was in. The figure below shows a boxplot of resort ticket pricing across different. Some numeric data looking at total state area, the population, the resorts per state, the total skiing area, total night skiing area, total days open, and state resort density. Also while exploring the data I found that states with a high density of resorts to the size of the state as well as a high population is an important factor to ticket price. Some features that seem correlated to price include that vertical drop, total skiable area, chairlifts, and fast quads. To handle different states' individual markets I feature engineered ratios of skiable area, days open, terrain parks, and night skiing. Finally I added features that can show the ratio of chairs to the size of the mountain.

Modeling:

I preprocessed and trained 3 different models. First I created a baseline model by just taking the average, this performed poorly. I then trained a linear model with a 70/30 train test split that performed better with a Mean Absolute Error of around 11. This may though have been due to overfitting, so I then used a Kfold strategy to prevent this and found a more durable model. In this linear regression model vertical drop was by far the most important feature.

The Winning Model:

The third and best model I then used was a random forest model, and after some hyperparameter tuning came to a MAE of 9.58 which was the fewest absolute error and is why I will be using it. Its most valuable features were fast quads and runs. Below is a histogram of feature importances. 

Conclusion and Recomendation:

Currently Big Mountain Resorts ticket price is $81. The predicted model price is $96.55. Approaching business leadership I would show that for the current amenities offered by the mountain they should be charging more. Additionally I created a scenario of adding 1 run, 1 chairlift, and an increase in vertical drop by 150ft. After these changes I found this could support a further price increase of $10.36 creating and increase of expected revenue of $18,134,058. As far as run closers go, I would suggest starting with 1 as I find no price drop off in price after 1 closer. I found a large $1.75 decrease after 5 and would not recommend closing more than that.

Further Thoughts:

One deficiency of the data was there was little information on the quantity of guests visiting resorts each season. I expect a higher ticket price with the same amount of amenities will decrease the number of visitors. There is also no data on operating costs of lifts, snowmaking, and night skiing lighting that could prove useful. Business leaders may be surprised by the higher modeled ticket price because they are setting a price based off of competitors, when they have a superior mountain in terms of features. To test the model and different mountain features I would implore them to use the predict increase function to test and explore.