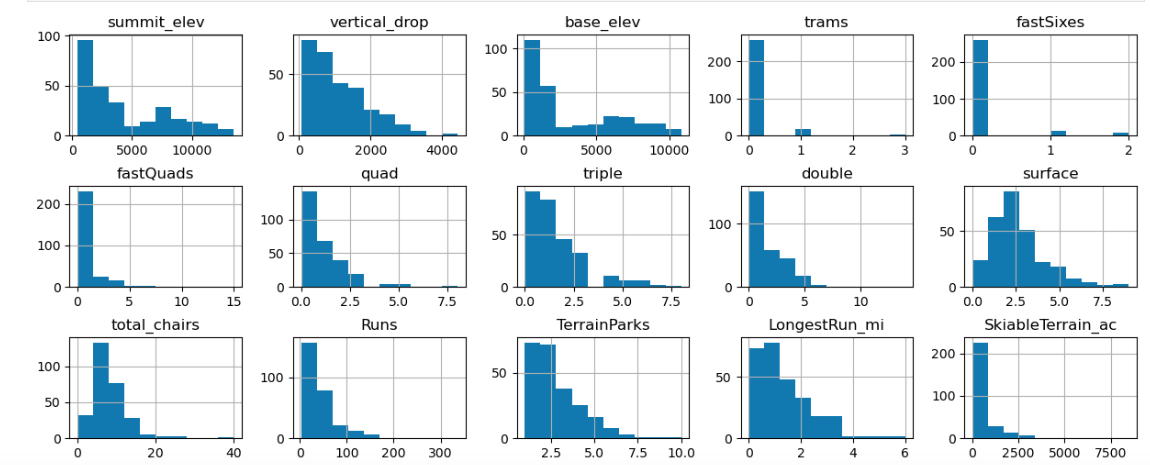
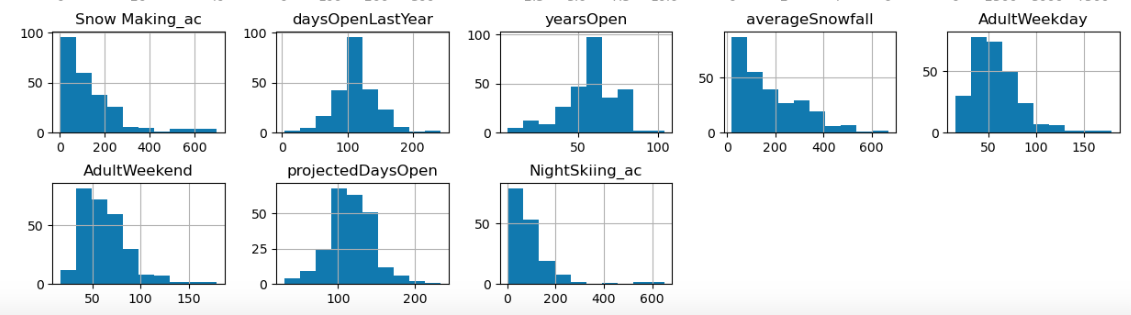
Bear Mountain Resort in Montana is looking to improve their investment strategy or cut costs to fund $1.54M to maintain a new lift that they plan to install. They have provided us with a .csv file that has data for most resorts across the country. They would like to see a model which predicts the best strategy for them. The data was all in .csv file, so the model had to be built from scratch which meant cleaning the data, then exploring and analyzing it, and then building the model to predict possible price increase.

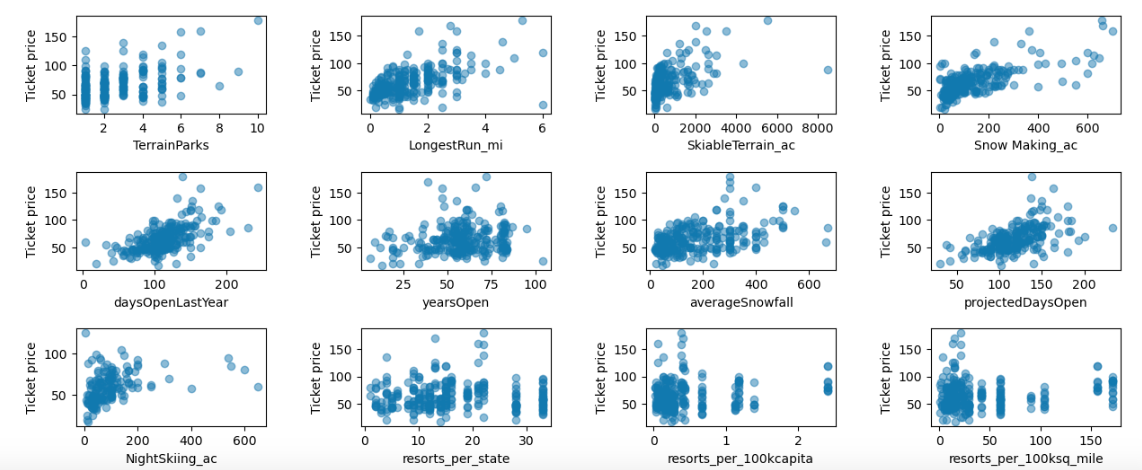
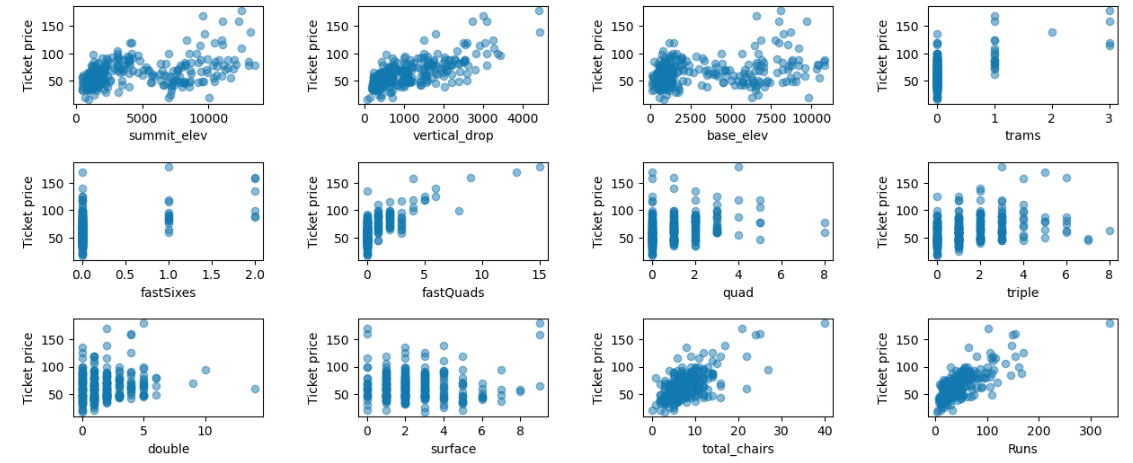
Data Wrangling:

The data in the .csv was quite well organized with all columns labeled and most data available. After we cleaned the data, filled in the null values and dropped columns that were not relevant, the below shows the spread of data in the csv. The .csv did not mention the number of visitors the resorts got every season, so we created another dataframe with sate population.

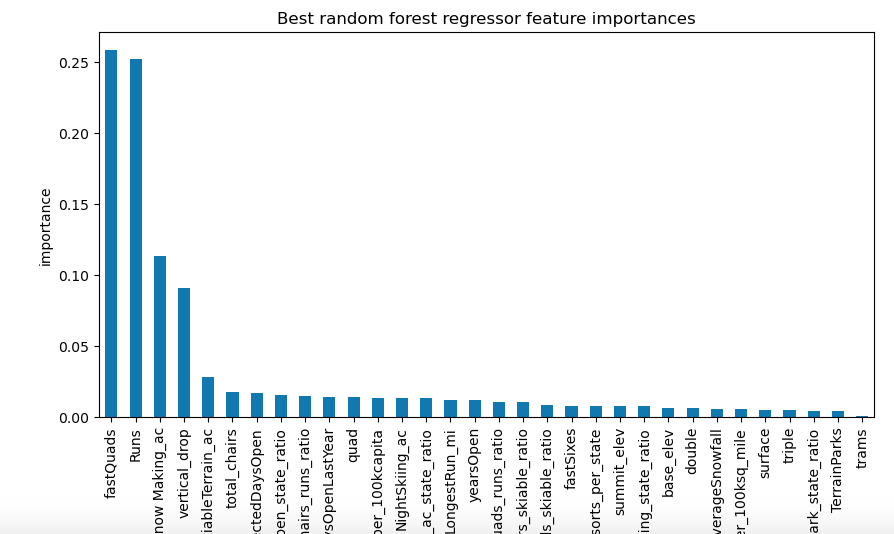




In the EDA phase, we decided to treat all states with equally. We did a PCA scaling and transformation and saw that a couple of features pushed a few states to be outliers but no other pattern was established. We derived a few other features like ski area per resort vs state etc and created a scatter plot to see that a few features like vertical\_drop, fastQuads, Runs, and total\_chairs were highly correlated with the ticket prices.



At this point we moved on to training and testing the model we were going to build. We used Dummy Regressor, Linear regression and random forests and cross validated using GridSearch. This way we built and tested the data on our training data and did not touch the test data till we picked a model. The R2 and other metrics showed almost a dollar difference in favor of random forests and we decided to go with it. This also gives us the best features we can used that will build the most accurate model for predictions.

Next, we compared the top ten features to see how Bear Mountain compares with the other resorts. The features that came up as important throughout the analysis were: vertical\_drop, Snow Making\_ac, total\_chairs, fastQuads, Runs, LongestRun\_mi, trams, SkiableTerrain\_ac

We saw that Big Mountain is a leader in Vertical drop, snow making area, number of chairs, fast quads, runs, longest run and skiable area. We then used out model to see if there is a scope for price increase given these facts and what revenue would be generated because of the price increase.

The winner was Big Mountain is adding a run, increasing the vertical drop by 150 feet, and installing an additional chair lift

[35]:

This allowed for a $1.99 per ticket price increase and that generated over $3.5M in revenue which is more than enough to offset the operating costs of the new lift.