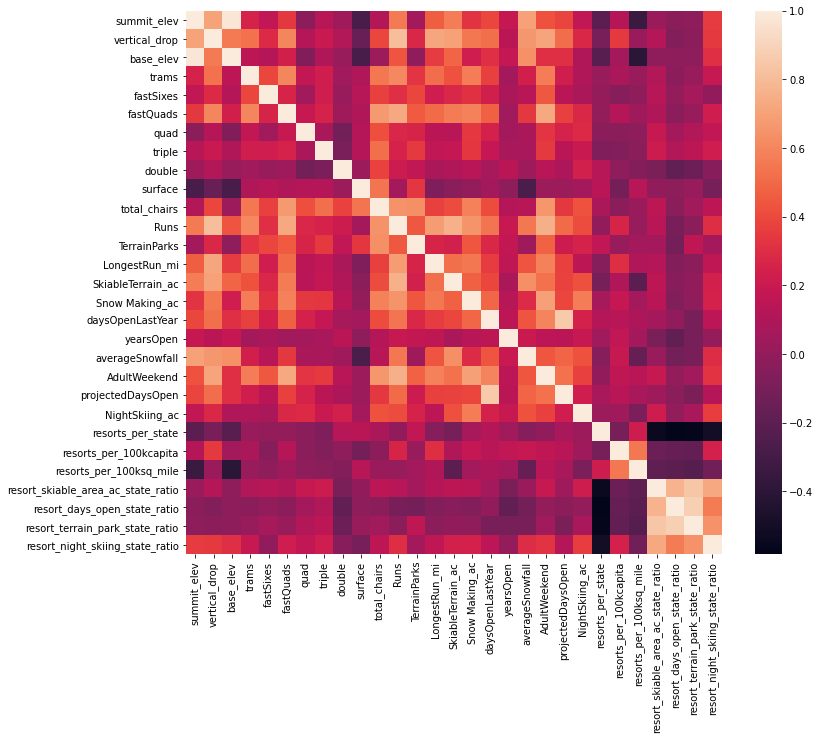
**Big Mountain Resort Ticket Pricing**

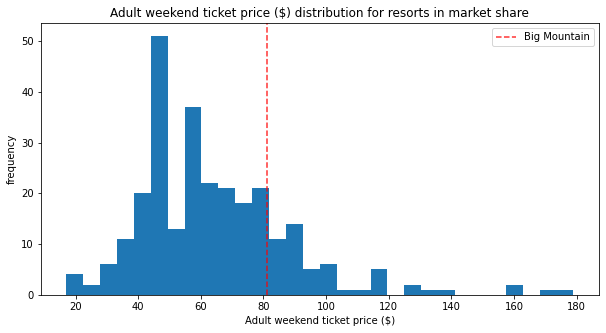
The Dataset given to us contains 330 resorts from all over the country and gives us their different facilities like skiing area, night skiing, year open, vertical heights and many more. These features or facilities are used to predict the ticket pricing for our resorts for this upcoming season. We have also used the population and area of the states into account to create different parameters like resort per state, resort per 100k Population, resort per 100k sq mile area etc to make prediction of pricing clearer.

There are only 35 States that have skiing resorts as per the dataset. We have made the pricing for tickets based upon the weekends prices of the different resorts as we have more null values in the weekday prices in the data. Ticket price could climb with the number of resorts serving a population because it indicates a popular area for skiing with plenty of demand. The lower ticket price when fewer resorts serve a population may similarly be because it's a less popular state for skiing. The high price for some resorts when resorts are rare (relative to the population size) may indicate areas where a small number of resorts can benefit from a monopoly effect.

Firstly, we tried the baseline model which is average price for tickets by using Dummy Regression using mean as strategy then to test how close the number is to actual values. We use R-squared, mean Absolute Error and mean Squared Error. We get 0.0 o training set as expected and get 0.0031137940436898326 you can expect performance on a test set to be slightly worse than on the training set. Mean absolute error is arguably the most intuitive of all the metrics, this essentially tells us that, on average, we might expect to be off by around dollar 19 if we guessed ticket price based on an average of known values. Then we try a Linear model to the training set. our simple linear regression model explains over 80 percent of the variance on the train set and over 70 percent on the test set. Using this model, then, on average you'd expect to estimate a ticket price within dollar 9 or so of the real price. Without using feature selection. With using feature selection our simple linear regression model explains over 79% of the variance on the train set and over 61% on the test set. Using this model, then, on average you'd expect to estimate a ticket price within dollar 10 or so of the real price.



For selecting the best fitting feature uses cross validation technique so we get around 8 main features that mostly affects the pricing out of these features.Random will be the best model because a random forest model has a lower cross-validation mean absolute error by almost $1. It also exhibits less variability. Verifying performance on the test set produces performance consistent with the cross-validation results.



Big Mountain Currently charges an adult ticket for 81 dollars. We have done some modelling based upon the data set of 330 resorts around the USA. We came with the increase in the pricing Big Mountain Resort modelled price is dollar 95.87, actual price is dollar 81.00. Even with the expected mean absolute error of dollar 10.39, this suggests there is room for an increase. We can reduce the run up to 10 so that our operational cost can came done and also increase the price of the ticket as we are adding the chair lift This scenario increases support for ticket price by $1.99 .Over the season, this could be expected to amount to dollar 3474638 over 350000 people visits our resort he additional operating cost of the new chair lift per ticket (on the basis of each visitor on average buying 5 day tickets) in the context of raising prices to cover this.