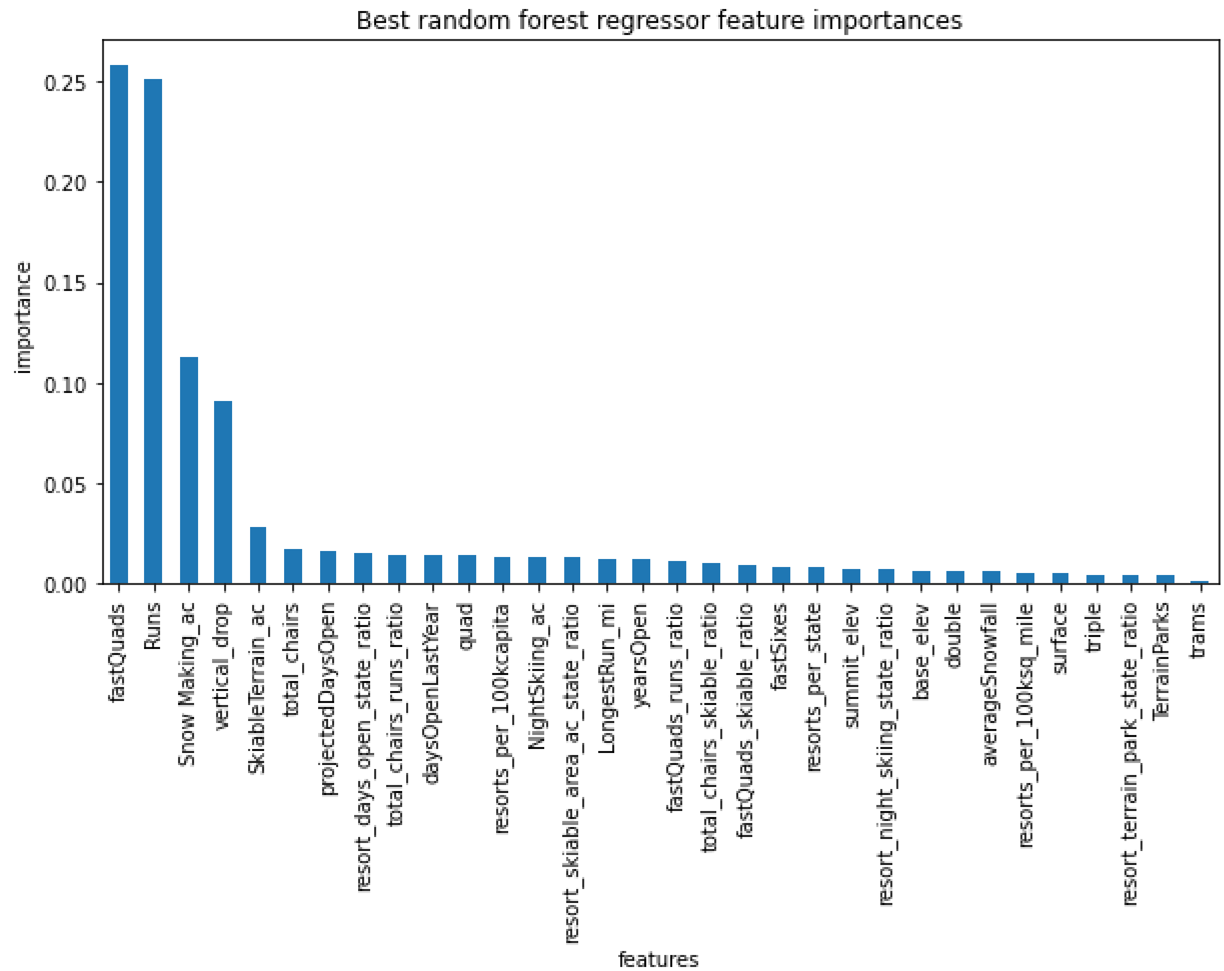
Big Mountain case study summary recommendations

This project is geared towards achieving a satisfactory solution to the question at hand, which is: How can the Big mountain Resort offer an increased seasonal resort demand even after an increased resort price, or how can profits be maximized by either cutting operational costs or resort incentives that will offer value that customers are willing to pay top dollars for? This relatively difficult question to answer was broken down and tackled through the data science methodology steps.

The problem was defined, the starting dataset was wrangled and cleaned to obtain the proper format, an EDA (exploratory data analysis) was performed to determine if there are any critical relationships between our target variable ***AdultWeekendprice*** ticket and the other column variables of the ski_data dataset. In addition, the working dataset was preprocessed so as to prepare it for model training, this including imputing any missing values or encoding any categorical variables that may exist. The data was then trained with different models, in our case, we tried out a simple linear regression model and a random forrest for both training and testing on our train-test data split.

For the linear regression model, the top features to consider that impact the Adultweekendprice ticket were: *vertical_drop, Snow Making_ac, total_chairs, and fastQuads*. For the random forrest regressor, the top features to consider were the *fastQuads, Runs, vertical_drop, and Snow Making_ac*.



Due to the lower mean absolute error value (~9.53 vs 11.79 for linear regression) obtained during a *5-fold cross validation* for the random forrest regressor, it was concluded that the random forrest should perform better in modelling our dataset than a linear regression. In the modelling section of the data science method, the random forest model was organized and used to model our data in a model pipeline which was previously defined.

From the model ticket price predictions, the random forrest regressor predicted a ticket price of $95.87 while the actual ticket price is $81.00, which suggests that either the Big mountain ski resort is either underpricing it's resort tickets in relation to what other resorts are charging in the market. It is important to look closely at the resort features and how they influence the model predictions. In the modelling step, a ticket price increase prediction custom function was also defined for being able to predict a ticket price increase from variations in key features values/combinations.

Based on the assumption that the resort should expect 350,000 persons and that each person coming to the big Mountain resort will purchase 5 day tickets, the expected revenue for the resort was seen to amount to 3,474,638 dollars, which would require a ticket price increase by 1.99 dollars ~ approx. 82.99 dollars. The price increase comes from the fact that the key features of Runs, vertical_drop, total_chairs, and Snow making_ac all increased by an additional 1 run, 150ft in vertical drop, 1 more chair lift, and 2 acres of land covered by snow making machines. The same predict_increase function call which neglected Snow making_ac as a feature resulted in the same ticket price increase of 1.99 dollars, which informs me that the ticket price is not too sensitive to incremental increases in features or in the using Snow making_ac as a key feature.

Further investigation into different combinations of key features should be looked into to see if the ticket price can be increased even when operation costs from additional chair lifts are added to the resort, which will help in maximizing profit while not losing demand from reduced resort facility utilities. I also suggest additional columns for resort operational cost metric be added so the dollar-value for operational costs can be quantified accordingly.