Big Mount: Increase Ticket Price or Cut Cost

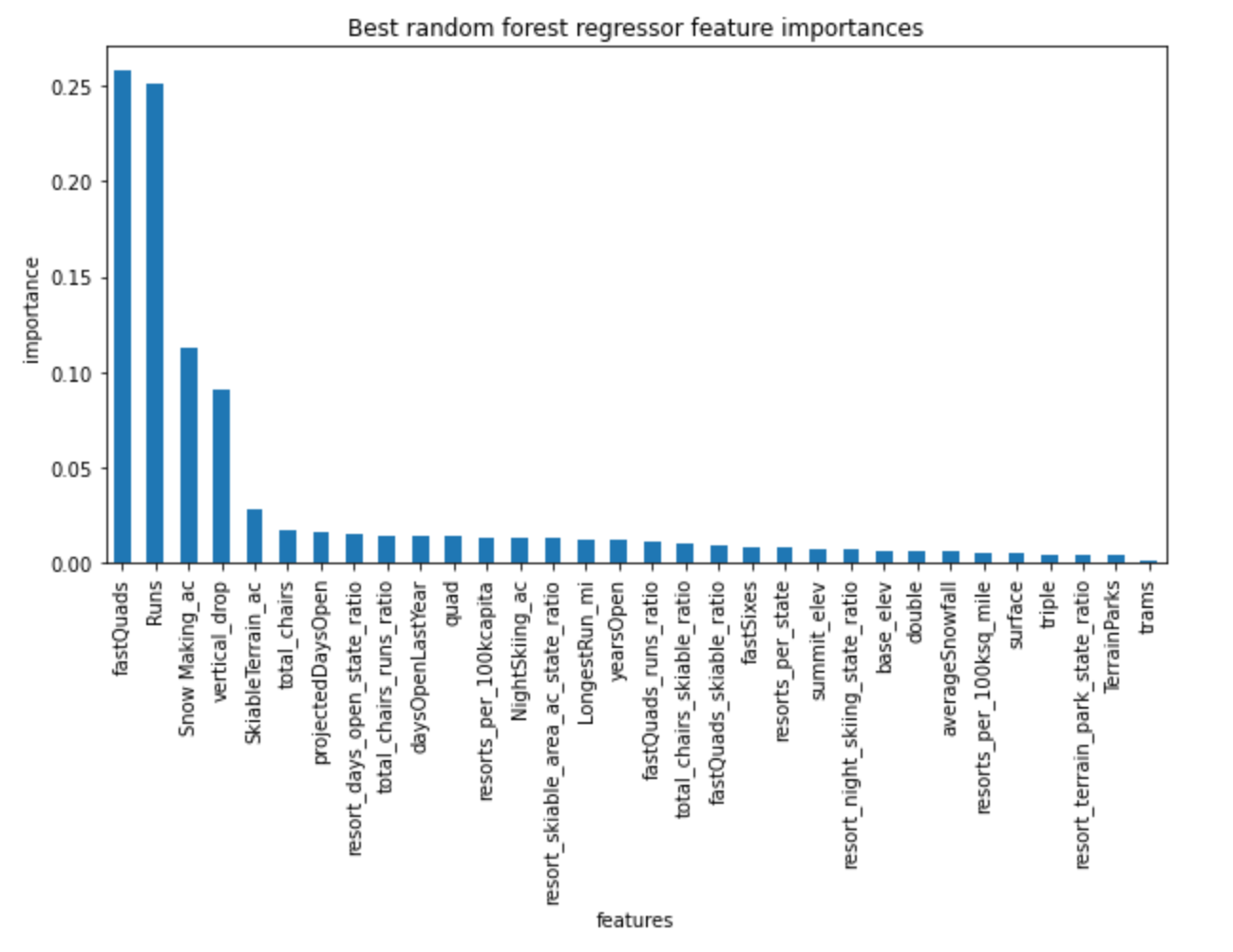
Big Mount must decide before the start of next ski season if they can increase their ticket price to add an extra lift or cut cost for maintaining all the facilities.

Important features for ticket price prediction:

To batter predict the ticket price for Big Mount, the weekend adult ticket price is used to fit a Random Forest model. Big Mountain currently charges $81.00, but the model predicted price is $95.87. Although Big Mount already has facilities at he higher end compared with all the other resorts, their ticket price is among the highest in Montana.

The model found that “fastquads”, “runs”, “snow making\_ac”, and “verticle\_drop” are the most important features that affect the ticket prices (shown in red box in Figure 1).

Figure 1. Best random forest repressor feature importance.



Recommended scenarios including:

1. Adding a run, increasing the vertical drop by 150 feet, and installing an additional chair lift increases support for ticket price by $1.99. Over the season, this could be expected to amount to $3474638.
2. Big Mount can also try out gradually closing 3 to 6 runs to test out the loss in tickets sold vs. cost reduction.

Model could be more accurate if information of facility costs could be provide, strategies may be plausible, including:

1. Reducing more acres of snow making may make a difference if cost of snow making per acre was available.
2. Night-time skiing cost could be reduced without significant reduction in ticket price.
3. Add more fastquads may also be a good reason to increase ticket price given the cost of adding this facility.

Data wrangling, EDA, and preprocessing before the final modeling:

ski\_data is the original data that had 330 rows X 27 columns, with each row as a unique observation. Big Mount data has two ticket prices: weekdays and weekend. ‘AdultWeekdays’ column was dropped because it has more missing value than ‘AdultWeekEnd’. Rows missing ‘AdultWeekEnd’ price values and column with mostly missing data (‘FastEight’) were removed to a final of 277 rows X 25 columns (‘Name’, ‘state’, ‘Region’ are categorical, all other features are numerical). State information with population and total area were used for generating new values that may be important for predict the ticket price (277 rows X 36 columns).

Heatmap showed the correlations between AdultWeekend ticket price are positively related with “verticle\_drop”, “fastquads”, “total\_chairs”, “runs” and “snow making\_ac” (Figure 2. Shown in green box).

Data was then split into 70% train and 30% test sets. Random Forest and simple linear regression models were compared and Random Forest was finally chosen according to the cross\_validate results: The Random forest regression meet the estimates, mean MAE is 9.5 for test split Vs (9.6+-1.3) from train split. It is closer to the train estimate, compared with simple linear regression model: 11.8 Vs. (10.5+-1.6), and has less variation (1.3 for Random Forest model vs 1.6 for the simple model).

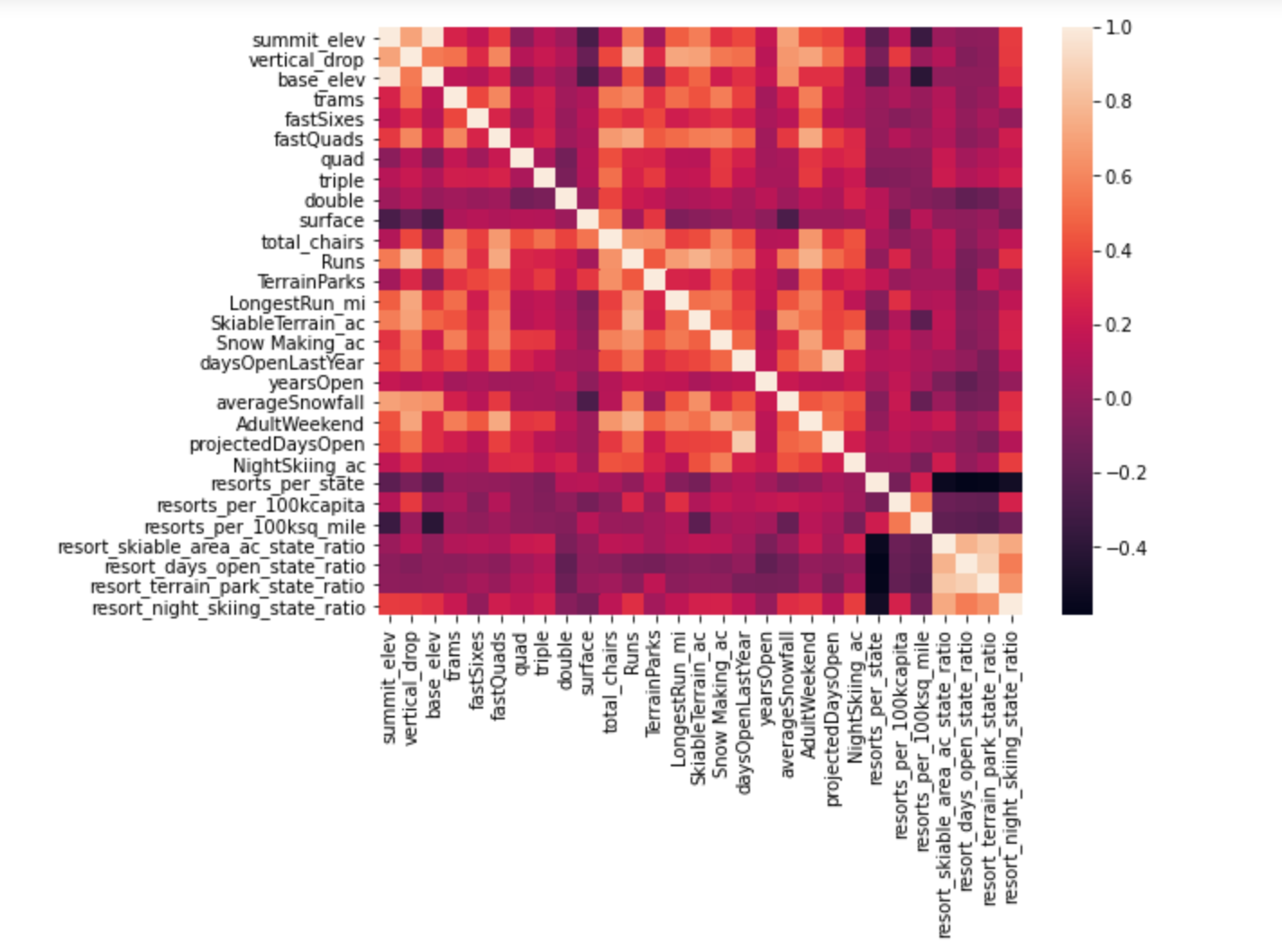


Figure 2. Correlations between numeric features