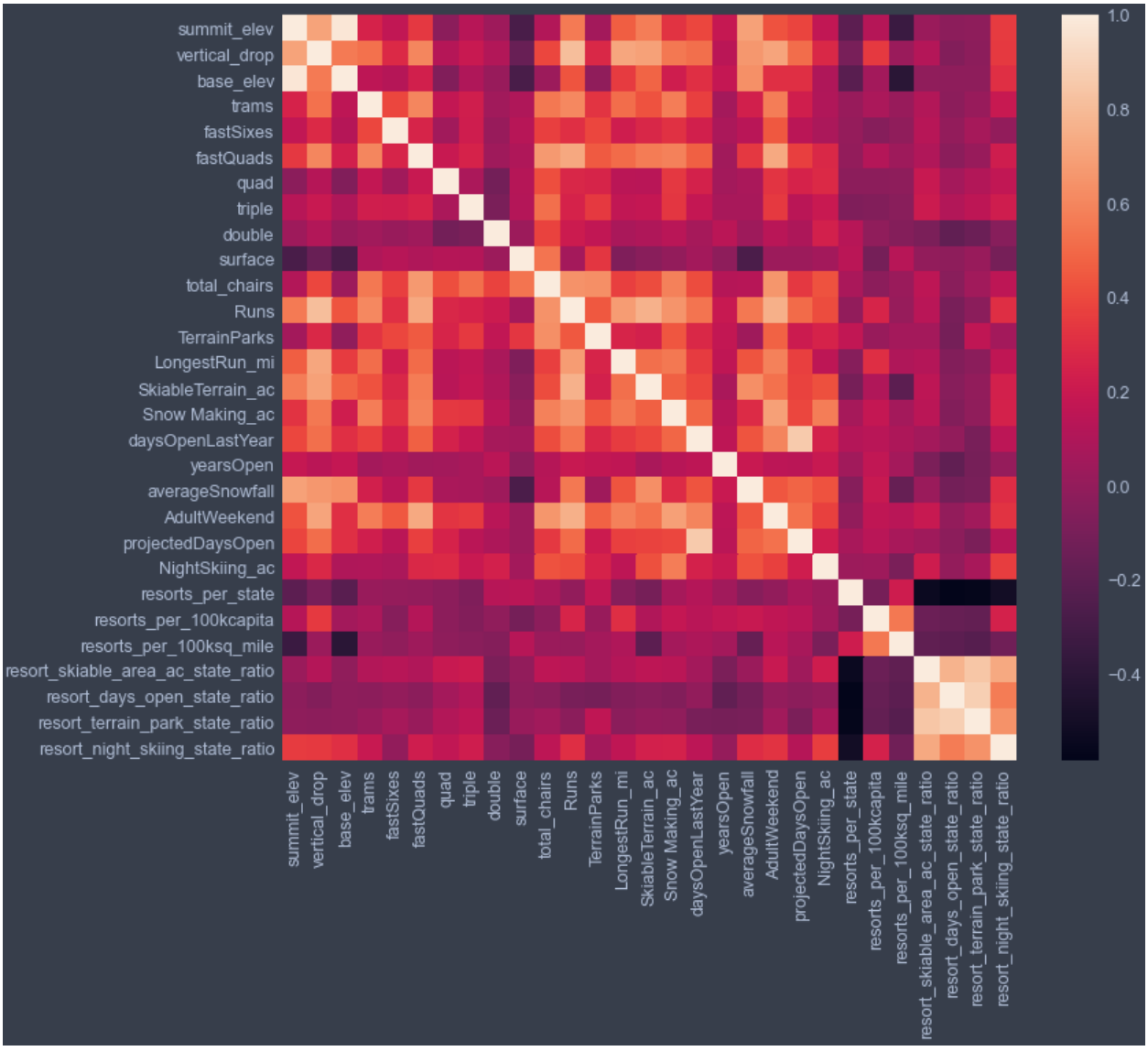
In this page, I will be going through a simple yet detailed summary of my analyst and recommendations for Big Mountain Resort based on the data set provided by Springboard's Guided Capstone.

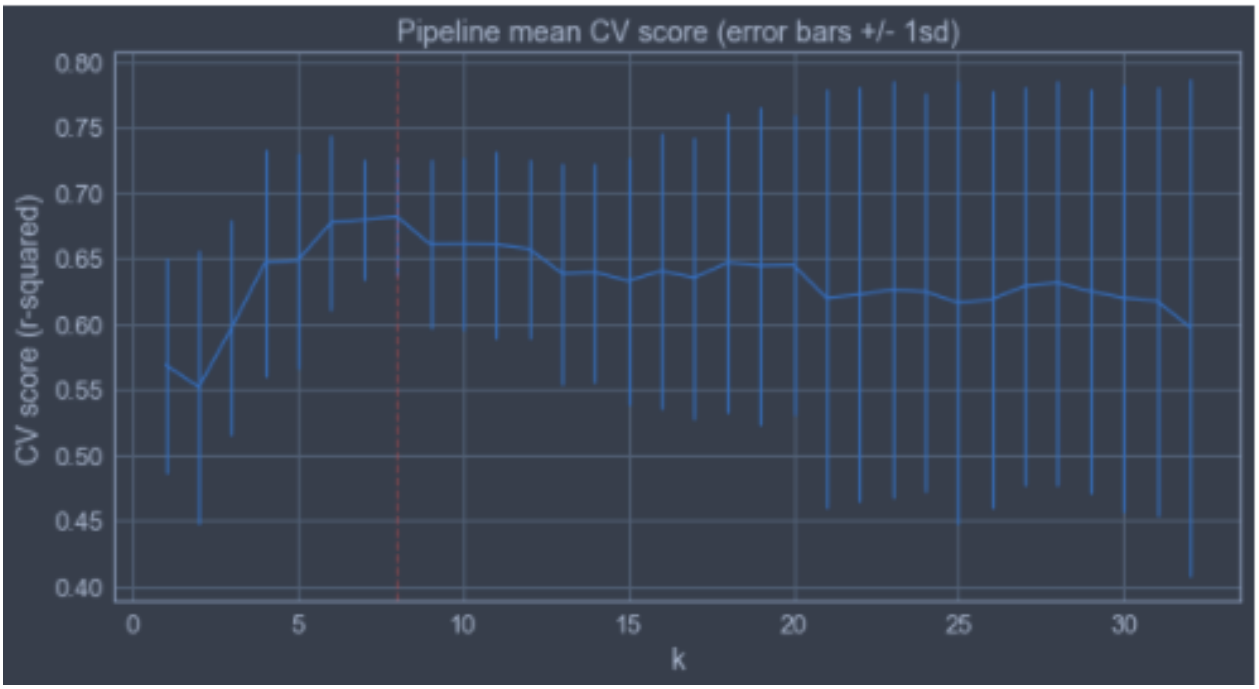
First and foremost which features are the most correlated? To answer that lets take a look at this feature correlation heatmap.



This may look very complicated but all we need to focus on are the relatively white boxes as this shows the highest correlation factor between the two corresponding features. This can be clearly seen in favor of the features located on the top left and bottom right corner. In this case, the summit and base elevation are quite highly correlated. As we turn our attention to the bottom right

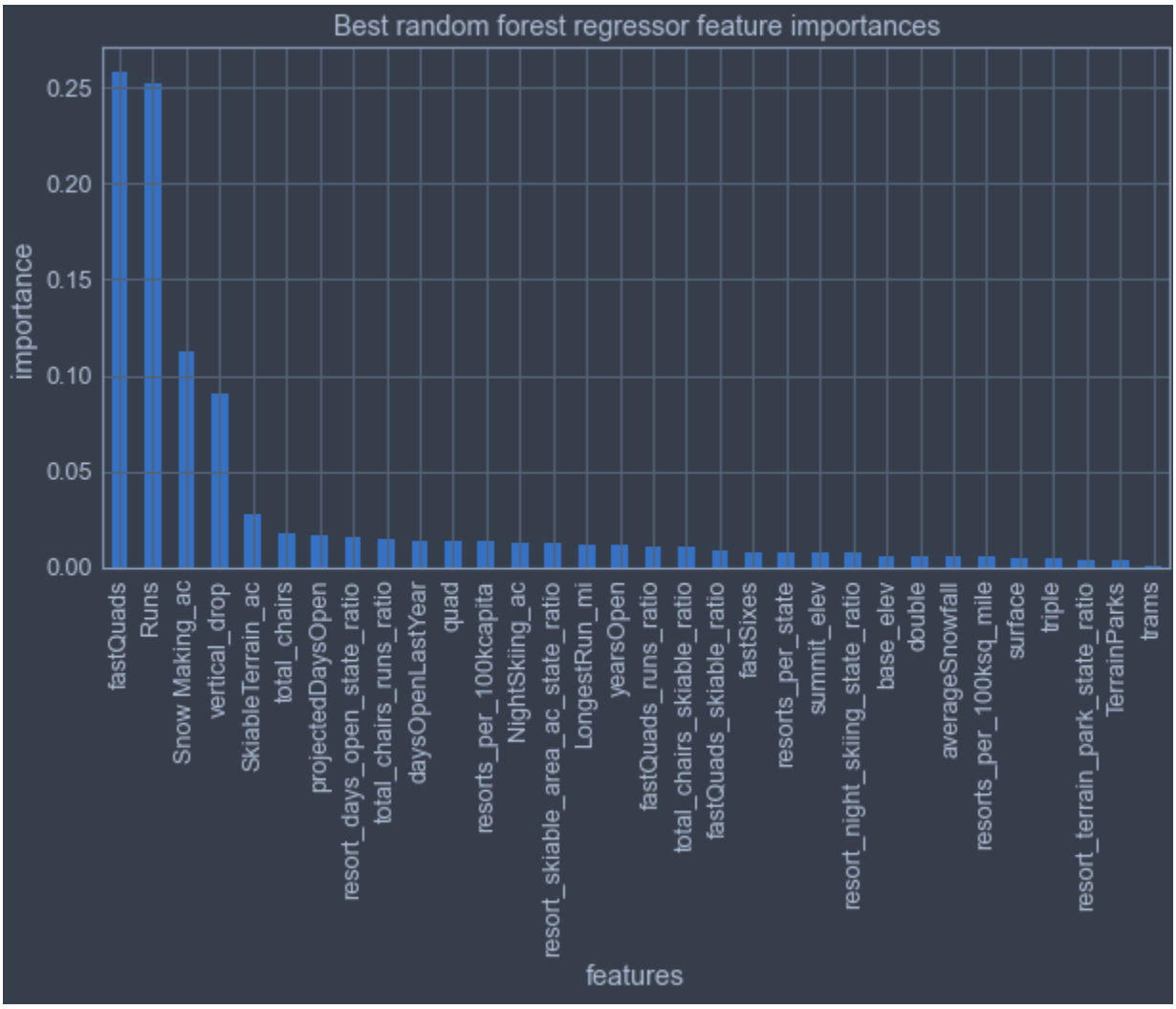
corner we noticed the introduced multicollinearity with a few selected ratio features. This is all good however also noticed the strip of dark lines that shows the negative correlation of the ratio features with the number of resorts in each state. This makes sense since the higher the number of resorts in a state, the lower the share each of the other state features will be. One important point to take note is the positive correlation between the ratios of night skiing area with the number of resorts per capita. This shows that the more densely located a resort is with population the more night skiing is provided.

Next I used an initial 70/30 ratio to train and test the model using linear regression model. I then assess performance using cross validation and ran a GridSearchCV hyperparameter to find the best k value (number of features selected or used to predict the model). As shown below, the CV score drops after k=8. By increasingly overfitting a model, we expect greater swings in performance due to different points moving in and out of the train/test folds.



Besides the linear regression model, I used the random forest model to investigate some different hyperparameters. The below figure shows the dominant top features that are in common between the random forest model and the linear model. By running both models, the mean absolute error of the linear model came out to about 11.79 while the random forest model was about 9.54. Since

the random forest model exhibits less variability and lower cross-validation mean absolute error, the random forest model was chosen.



Finally, we come to the modelling part of this project. I started by comparing weekend ticket prices of Big Mountain with all other resorts and only in the state of Montana. The current pricing of Big Mountain resort is at $81. When comparing with all other resorts, the modelled price is $95.87 with a mean absolute error of $10.39 which suggests there is room to increase ticket price. However, at the current price point, Big Mountain is charging the highest in Montana state. By comparing further on the other surrounding features, it was found that although Big Mountain only has a slightly above average vertical drop, it has a significantly high snow making area, one of the largest skiable terrain with a high number of runs and includes one of the longest run amongst all the other resorts. In terms of services features, Big Mountain has 3 fast quads which puts it

high up that league table while still having amongst the highest number of total chairs excluding outliers.

By further analyst on the different plausible scenarios provided by the company, it was found that closing one run makes no difference, while closing 3 to 5 runs incurs the same losses, and beyond 5 runs would decrease support for ticket price and revenue significantly. Comparing scenarios the rest of the scenarios: adding a run, increasing vertical drop by 150 ft, installing addition chair lift, adding 2 acres of snow making, or increasing longest run by 0.2 miles and guaranteeing snow coverage by adding 4 acres of snow making capability, the results were predicted to be negligible which suggest that we could ignore trying to improve on these features.

The additional operator cost of the new chair lift per ticket is about $1 per ticket. For future improvements, I would recommend modeling pricing based on Montana and its closest neighbouring state to get a better view of the business overview. If business want to test and progress with run closures, they could group runs or courses by special events that will allow for test period without affecting the everyday customers. By conducting special events or classifying specific courses to be, there could be potential for limited tickets with premium pricing.