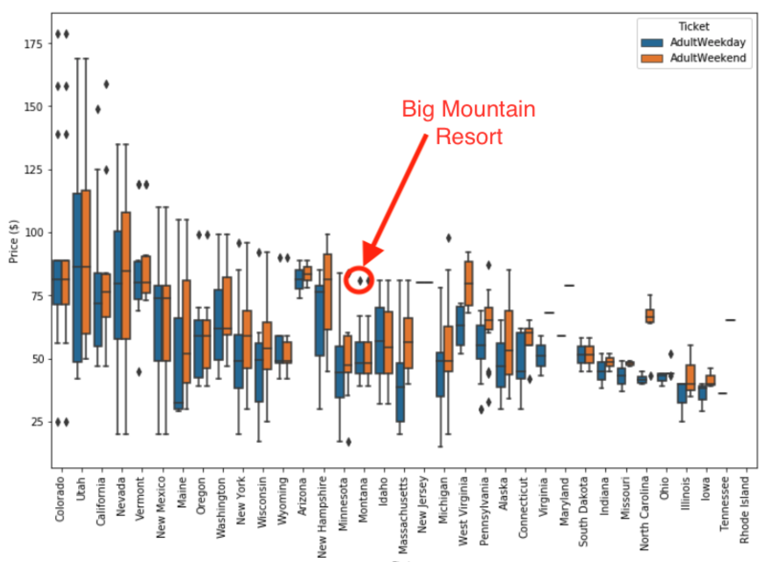
Data Analysis for Big Mountain Ski Resort

Problem Statement:
What opportunities exist for the ski resort, Big Mountain Resort, to increase revenue to account for the increased burden of $1.54M in added maintenance costs within the next year?

Scope:
Limited to ticket price adjustments to impact revenue.

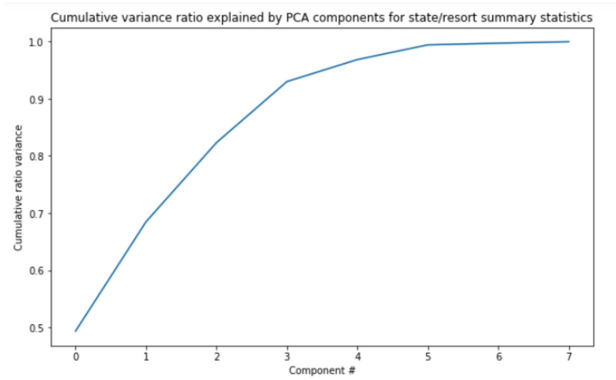
Constraints:
Impacted by last year project which increase costs by $1.54M which must be accounted for this year. Big Mountain is already premium price for ticket at $81 for adult, which is highest for Montana.

Original data imported as CSV file with shape (330, 27).
A comparison of data values on initial observation indicated that for Big Mountain Ski Resort, out of 330 data points it was 12th best for longest run distance, 5th highest in Skiable acres, 8th largest in Snow Making acres, 20th best for number of Runs, and 52nd highest in Adult Weekend ticket price. Montana as a state ranks 10th on resort density per state. The ticket price is already premium, but the resort has advantages and size that few competitors can match.

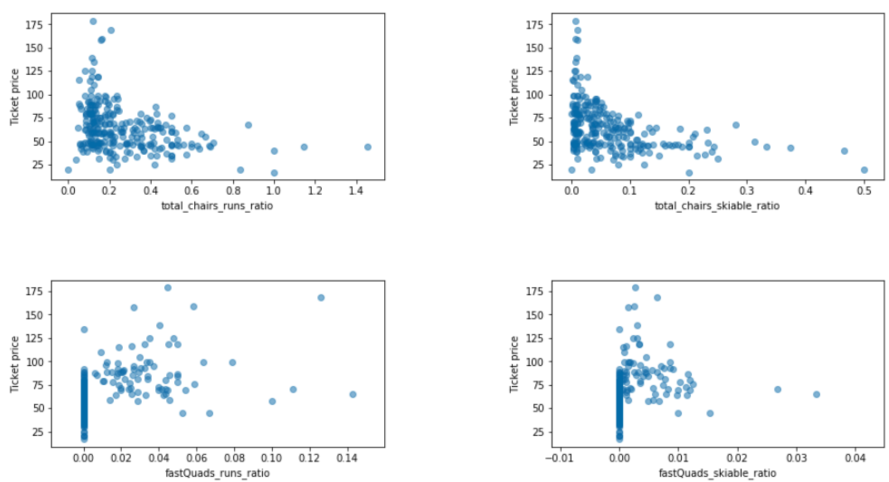


Data Wrangling:
From the initial data load, the data was reviewed and verified to not have duplication. Extreme values were researched for data issues and corrected. NaN values were removed and the FastEight category was removed due to NaN issues. Joined State population data to result in going from Orig Data rows 330, to cleaned 277 unique.

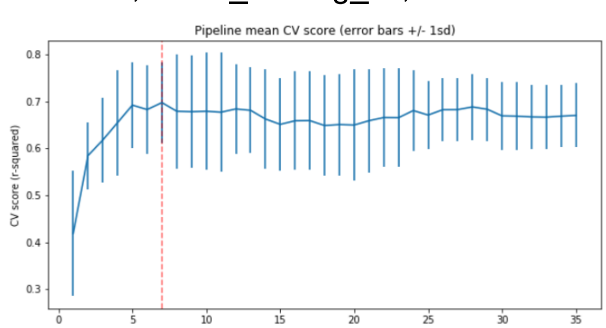
Exploratory Data Analysis:
Applied Principle Components Analysis (PCA) to scale data, fit PCA transform, and look at cumulative variance. Per graph, over 75% variance explained by 1st two components, and over 95% variance explained by 1st four components



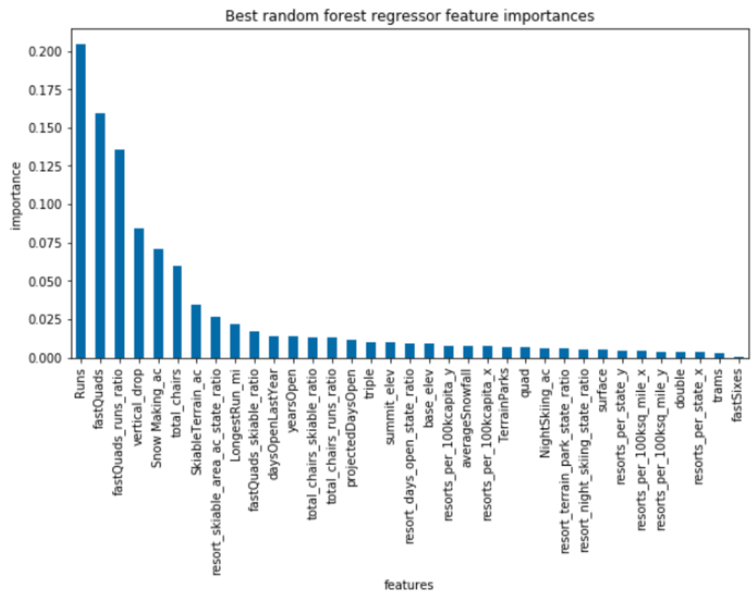
Scatterplots of features showed those with high probably relationships.
Strong positive correlations found for features: Total_chairs_runs_ratio, Total_chairs_skiable_ratio, fastQuads_runs_ratio, fastQuads_skiable_ratio.
Other features did not initially show significant correlation to the output.



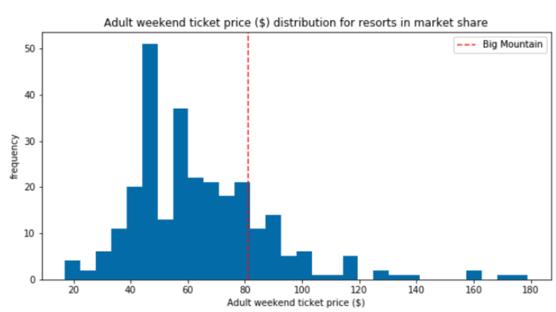
Preprocessing Data:
Used SelectBest to determine best value for K, # of subsets to slice data and cross-validation to estimate model performance. Determined k=7 best for model
Predicted top features impacting model were as follows: Vertical Drop, total_chairs, fastQuads, Snow_Making_ac, Number of Runs.



Used sklearn's Random Forest Regressor (RFR) to perform fit as part of Cross-Validation. Tried both mean and median for hyper parameters
Predicted top four features to model were Runs, Fast Quads, Vertical Drop, Total Chairs, and Snow Making acres. Did an MSE Comparison: Linear Regression 301.12 but RFR 190.2 Ultimately Choose to go forward with RFR for lowest variation.

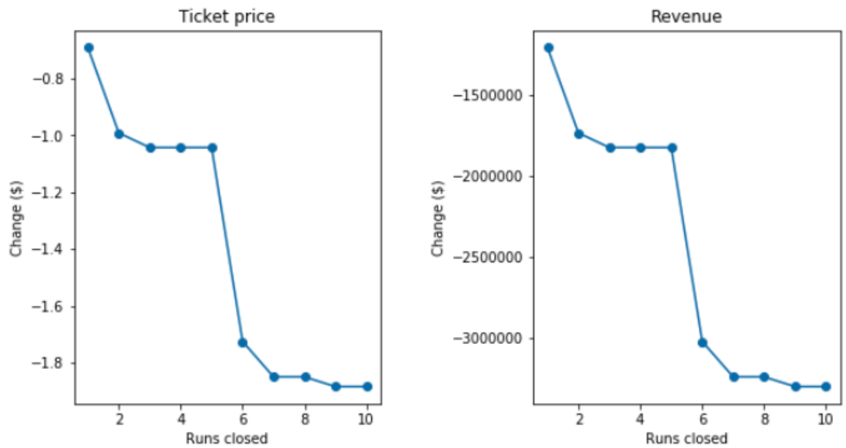


Modeling:
Loaded data into the model and prepared for business scenarios. Removed Big Mountain data and refit data (276 rows). Cross-Validated model (mean 14.01, std 8.19)
Created subplot for all important features. Looked at the placement of our outcome value: adult ticket price. At $81M is above average and highest for Montana
Reviewed the predicted top four features that impact value for ticket
Factors determined lower impact: trams, and skiable terrain



Business Scenarios:
1) Close up to 10 of the least used runs
2) Add a run, increase a vertical drop by 150 feet, add a chair lift
3) Add a run, increase a vertical drop by 150 feet, add a chair lift, and add 2 acres of snow making
4) Increasing the longest run by .2 miles, guarantee snow coverage by adding 4 acres of snow making capability.

Model Results:



1) If any combination or runs removed, ticket prices lower and revenue is lost.
2) Best model output prediction, increase price $1 ticket, increase revenue $2M/year
3) Same prediction as #2: increase price $1 ticket, increase revenue $2M/year
4) No impact to ticket price, no impact to revenue gain

Conclusion/Recommendation:
The Random Forest Regressor model was identified to have the best accuracy. Fitting the data for resorts comparable to Big Mountain, and merging population and size information provided the best fit model. All business cases applied to model

Of the 4 Business cases considered: Case #2 provided the best projection for ticket price adjustment. The increase of $1 to ticket price projects revenue gain by $2M which achieves the Problem Statement.

Note: Case #3 produced same ticket price justification, but required more capital costs

Model recommends to apply Business Case #2