The current pricing structure for ski tickets at Big Mountain Ski Resort is to take an average of all ski resorts of different sizes across all states. A preliminary examination of different ski resorts in different states and regions with different levels of population and resort amenities revealed a lot of ambiguity in the data that suggested revising resort amenity parameters to improve Big Mountain Ski Resort revenue. Due to a lot of noise in the data regarding state size in acres or state size in population it appears that new scaling columns would be useful to standardized per 100K of state population in the State Summary table: 'resorts\_per\_100kcapita' and 'resorts\_per\_100ksq\_mile' in the Exploratory Data Analysis phase.

Next an examination of the Cumulative Variance ratio was examined and the first two components accounted for 75% of the variance for this state level data. The first four of the eight components accounted for 95% of the variance.

Chart, line chart

Description automatically generated

These eight state level metrics were merged into the ski resort data and new scaled features were added to the ski resort data set. A pairwise correlation matrix was executed to view all pairs correlations of features as well as scatterplots of different features as the X variable versus ticket price as the Y variable to view the strength of their relationship.

Chart

Description automatically generated

In the scatterplot below It seems that the more chairs a resort has to move people around, relative to the number of runs, ticket price rapidly plummets and stays low. What we may be seeing here is an exclusive vs. mass market resort effect; if you don't have so many chairs, you can charge more for your tickets, although with fewer chairs you're inevitably going to be able to serve fewer visitors. Price per visitor is high but your number of visitors may be low.

Chart

Description automatically generated

As a baseline before training a machine learning model the average price of $63.81 was noted. Since many features measured were listed in many different units, with numbers that vary by orders of magnitude we scaled them to put them all on a consistent scale of mean zero and unit variance.

The select K best algorithm arrived at 8 coefficients which largely aligned with our correlation analysis in the EDA phase. Next, we looked at the Random Forest model and its top four coefficients for prediction coincided with the previous linear regression model: fastQuads, Runs, Snow Making\_ac, vertical\_drop.

When comparing the linear regression model the random forest model the random forest model has a lower cross-validation mean absolute error by almost $1. It also exhibited less variability. Verifying performance on the test set produces performance consistent with the cross-validation results.

In the final analysis, the predicted ticket price generated by the model is $95.87 and the actual ticket price is $81.

The business has shortlisted some options:

1. Permanently closing down up to 10 of the least used runs. This doesn't impact any other resort statistics.
2. Increase the vertical drop by adding a run to a point 150 feet lower down but requiring the installation of an additional chair lift to bring skiers back up, without additional snow making coverage
3. Same as number 2, but adding 2 acres of snow making coverage
4. Increase the longest run by 0.2 mile to boast 3.5 miles length, requiring an additional snow making coverage of 4 acres

Option 4 had no effect on revenue so it is not recommended. For option 1 the model says closing one run makes no difference. Closing 2 and 3 successively reduces support for ticket price and ergo revenue. If Big Mountain closes down 3 runs, it seems they may as well close down 4 or 5 as there's no further loss in ticket price. Increasing the ski run closures down to 6 or more leads to a large drop. Option 2 is the recommended route since it increases the ticket amount by $8.60 and option 3 begins to run into the Law of Diminishing returns.