Pricing Model Report - Big Mountain Resort

Problem Statement

Using a data-driven pricing strategy, how much of an increase in adult weekend ticket prices can the market support for Big Mountain Resort given what they currently offer, and what targeted operational investments and/or cutbacks would maximally support additional increases in ticket prices?

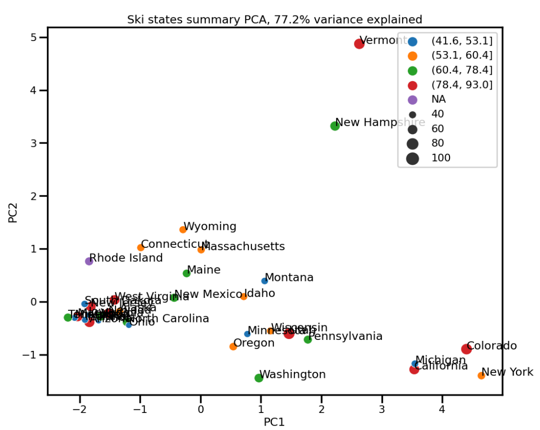
Note: *the above is revised from the original [problem statement](), as the problem itself has evolved with each step of this project (e.g., target feature of ticket prices narrowed to specifically adult weekend prices after initial data wrangling/exploration, omitted goal net revenue / specific timeline as it became clearer that the model would ultimately predict this for us).*

Data Wrangling

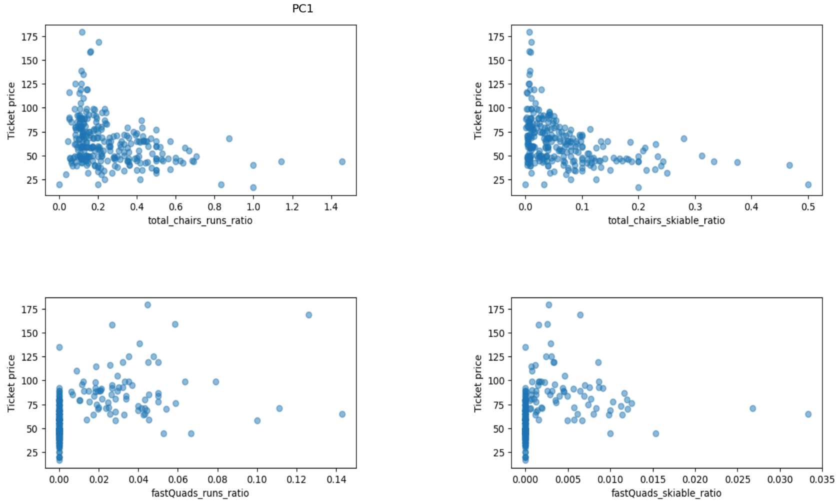
We received a dataset with multiple ski resorts in the region including ticket prices and resort features. After loading this data and verifying that Big Mountain's data was present without missing values, we performed an initial cleaning in which:

- ~ 7% of features (2 of 27 columns) were dropped, including
  - The "fastEight" feature, due to missing ~50% of values
  - The "AdultWeekday" ticket price, as "AdultWeekend" was chosen over it as the target feature due to a higher number of usable values (54 for "AdultWeekend" vs. 51 for "AdultWeekday").
- ~ 16% of resort records (53 of 330 rows) were dropped, including:
  - Two rows that had invalid values for the "yearsOpen" feature – i.e., values greater than 1000, likely representing erroneous entries of the calendar year a resort opened rather than number of years it has been open for.
  - Forty-seven rows lacking any ticket prices (both weekend and weekday).
  - Four rows lacking the chosen target feature of "AdultWeekend" prices
- The data were divided into categorical (State Summary Data) and numerical subsets.

Exploratory Data Analysis (EDA)



- We performed Principal Component Analysis (PCA) of the State Summary data, with a scatterplot for the first two components shown on the left.
- Resort density metrics appeared to have some influence on the second PCA component, particularly at the extremes (e.g., Vermont and New Hampshire had the highest resort densities), while states were fairly evenly spread across the first component.
- Ultimately, there was no clear relationship between state and ticket price, so it was decided that feature selection would consider all states equally - i.e., the State Summary data were merged back in with the numerical data.



Further visualizations revealed several features showing strong correlation with price (scatterplots, left). Most notable were

- inverse relationships with total chairs to runs ratio and total chairs to skiable area ratio
- potentially two different distributions of prices for having no fastQuads at all vs. having at least one

Preprocessing, Feature Engineering, and Modeling

- In general, we trained a series of **linear regression** and **random forest** models, using a simple mean as a baseline "model" for comparison.
- Preprocessing steps included splitting into training and test subsets, median imputation, scaling, and feature engineering.
- Evaluation metrics included R^2, Mean Absolute Error (MAE), Mean Squared Error (MSE), Root-Mean Squared Error (RMSE) – of most utility were [1] MAE, due to its ease of interpretation and ability to provide a sense of typical error magnitude, and [2] R2, as the default scoring method for 5-fold cross-validation.
- Initial feature selection was performed on the linear models using k=10 and k=15 with selectKbest, which identified top features such as Runs, vertical_drop, and fastQuads. However, these models performed worse than without such feature selection. Grid search with 5-fold cross validation revealed the ideal k-value to be 8, and provided an updated list of features, ranked by correlation coefficient (β).
- The same approach of grid search with cross-validation was then applied during training and hyperparameter tuning of the Random Forest models.
- Both types of models consistently ranked the same top four features: FastQuads, Runs, Snow Making_ac, and Vertical Drop. These suggest potential areas of investment or reduction for Big Mountain to increase their ticket value.

Winning Model & Pricing Recommendation

The Random Forest approach outperformed Linear Regression on various metrics, with a ~$1 lower error (MAE) in predicting ski resort prices. As such, the best RF model was chosen as the pricing model, generating the following recommendation:

> Raise the current adult weekend ticket price of $81.00 to $95.87 (MAE of ±$10.39) – an increase of $14.87.

Assuming 350,000 visitors and 5 tickets per visitor on average, this implies a potential revenue gain of ~$26 million.

Factoring in Big Mountain's recent investment in a new chair lift, an additional ~$1.99 increase would likely be supported; given the same assumptions,this would imply a potential gain of ~$3.48 million. However, this must be weighed against the increased operating costs to accurately assess ROI.

Further investments or feature reduction could support even further increases in the adult weekend ticket price; in particular, scenario modelling suggested this could be done with (1) additional vertical drop & chair lift, along with (2) closing a maximum of 1-2 least-used runs (i.e., closing > 2 runs will lower ticket prices).

Future Scope of Work

• **Additional data:** obtaining data on operating costs, seasonal staffing, capital investment/maintenance, and visitor satisfaction or demand elasticity may be helpful for more comprehensive modeling.

• **Making Sense of the Model Prediction:** Big Mountain's high ranking in facility metrics led to higher price predictions, which may surprise leadership if they feel constrained by local competition or internal budgeting. Presenting market comparisons and soliciting their perspective could provide further insight on the rationale behind the recommended pricing.

• **Model Implementation:** If the model is found useful, our team can package it into a user-friendly app or dashboard for Big Mountain's internal analysts to use, allowing modification of inputs (e.g., terrain, lifts) and real-time viewing of predicted price support/revenue.