Big Mountain Resort: Ticket Price Prediction and Scenario Modeling Report

Problem Statement

Big Mountain Resort wanted to find out if its current adult weekend ticket price reflects what the market would actually support, given the resort's facilities. The main goal of this project was to predict a fair market price using data from comparable ski resorts and to test how changes in amenities — such as adding a new chair lift, runs, or more snowmaking — might affect the ticket price. This information would help guide better pricing and future investment decisions.

Data Wrangling

The dataset included detailed numeric features describing each resort, such as summit elevation, base elevation, vertical drop, number of runs, chair lifts, fast quads, snowmaking area, and night skiing area. To better understand each resort's local market, I added population data from an external source. This allowed me to create new ratio features, such as the number of resorts per 100k people and resorts per 100k square mile, to measure competition and density. The data was cleaned by handling missing values, renaming columns for clarity.

The fastEight column was dropped because it had over 50% missing values, mostly zeroes. AdultWeekday was also removed due to a higher rate of missing values and the decision to use AdultWeekend as the target.

Rows missing both weekday and weekend prices (14% of original rows) were removed because they lack the target variable. Big Squaw Mountain was removed due to implausible value (yearsOpen = 2019).

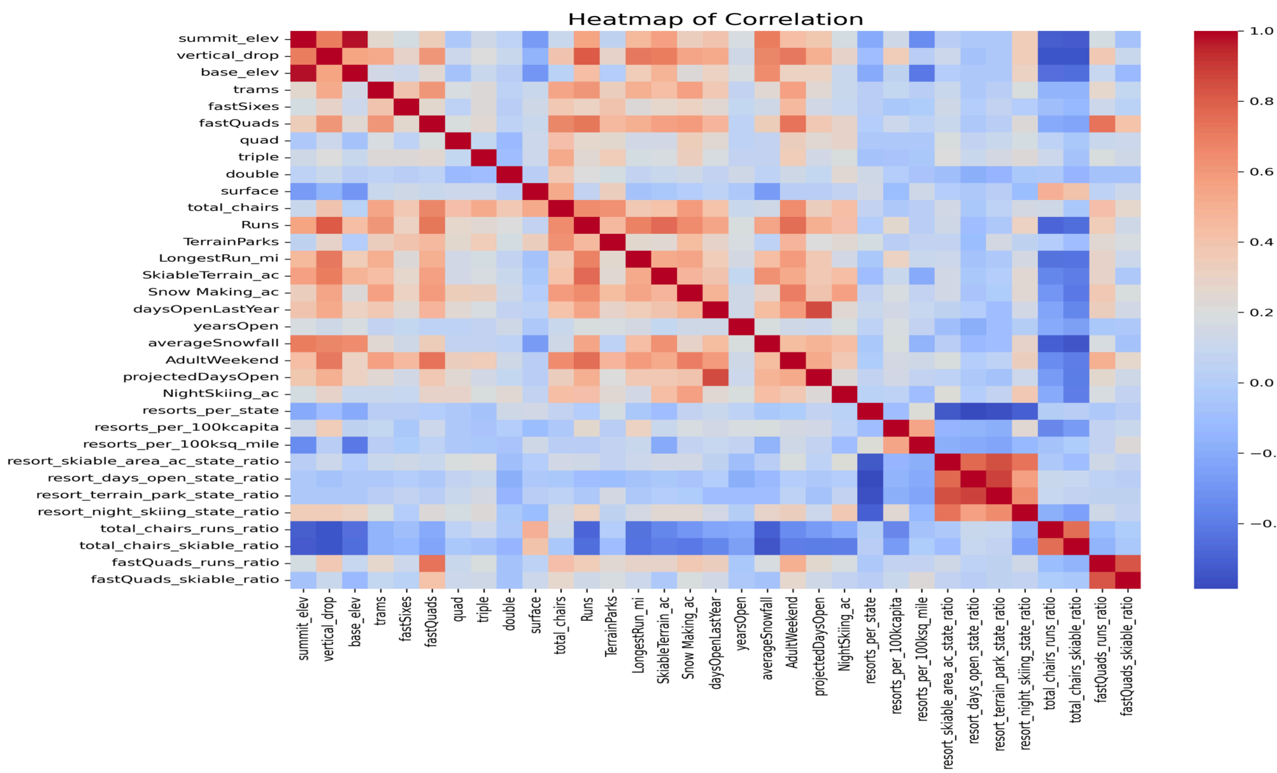
Corrected erroneous skiable area (SkiableTerrain_ac) for Silverton Mountain from 26819 to 1819 acres based on external source verification

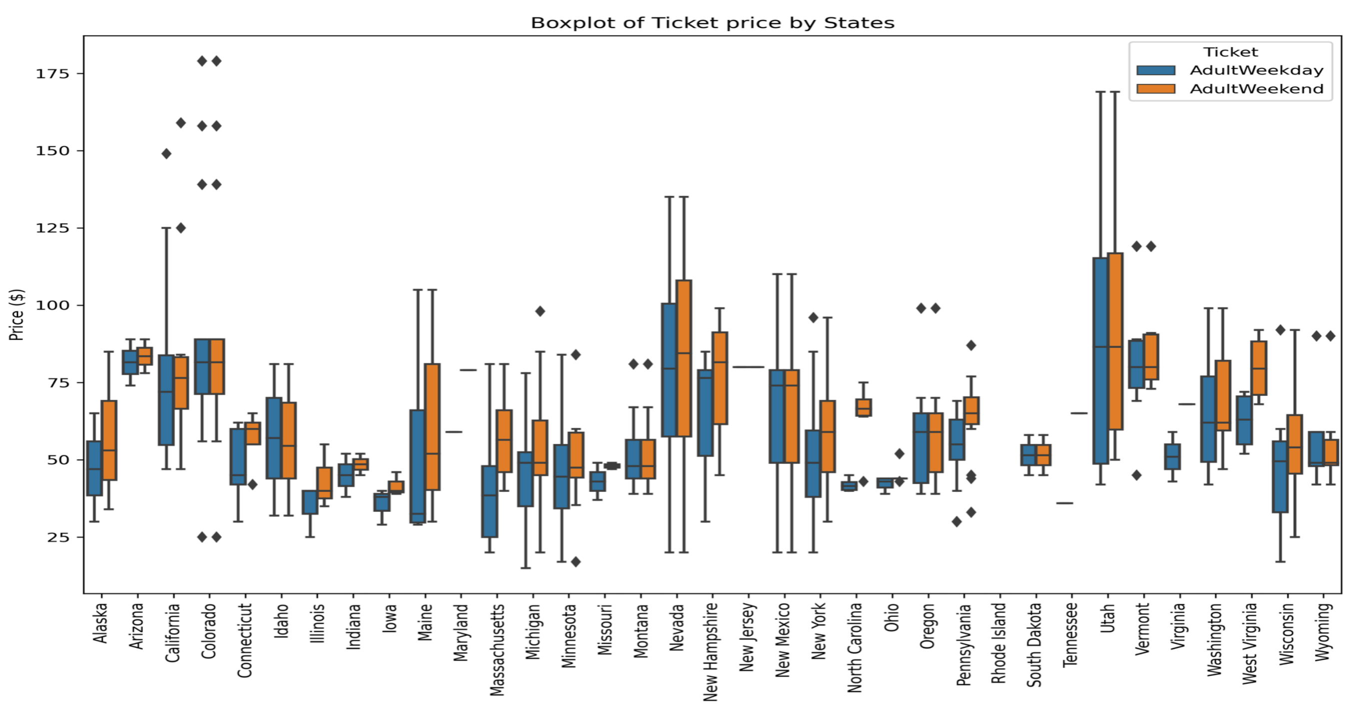
Exploratory data analysis (EDA)

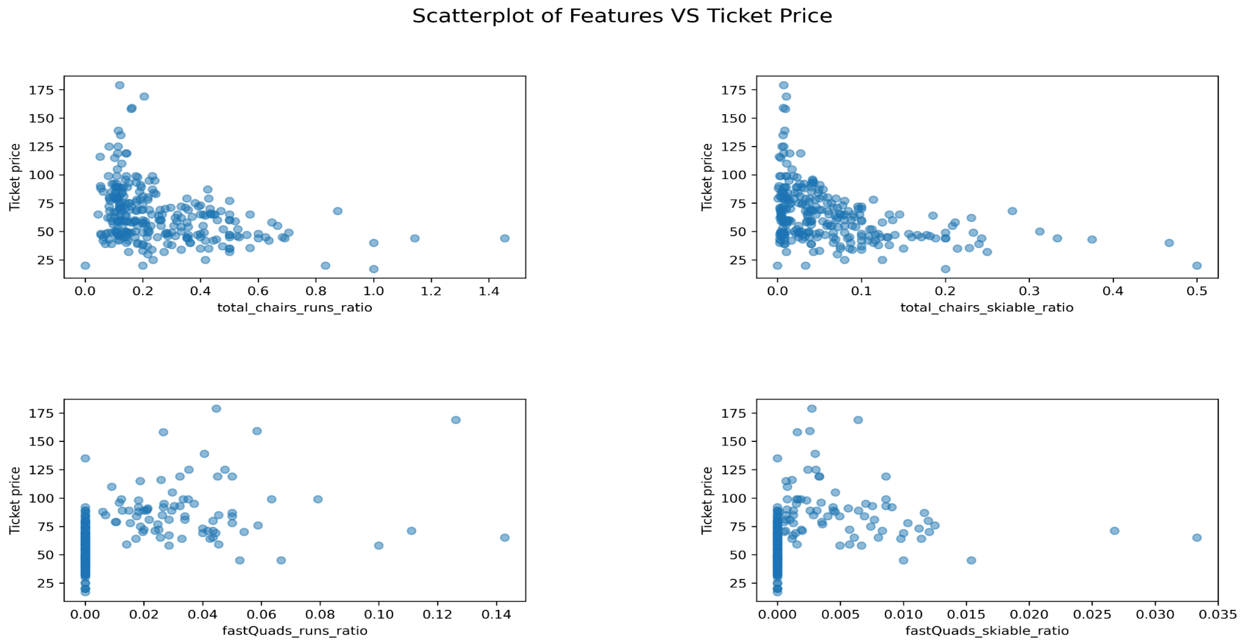
This exploratory analysis examined how resort features and state-level factors relate to adult weekend ticket prices. By comparing raw state data with engineered ratio features, we found that specific resort characteristics and each resort's share of its state's facilities were more useful for predicting prices than the state name alone. This guided our final choice of features for building an effective pricing model.

We used several visual tools to find patterns in the data. A correlation heatmap showed which numeric features were strongly related to ticket price and helped spot redundant columns. Scatter plots displayed how ticket prices change with key features like runs, lifts, and vertical drop, while a box plot summarized how ticket prices vary across all resorts. Together, these visuals made it clear which features influence pricing the most.

<u>Below are the figures the showing the Heatmap, Boxplot, and scatterplot :</u>







We also explored state-level factors like total area, population, total skiable area, number of resorts, night skiing area, and average days open. For example, Montana has a lot of skiable land but fewer resorts, while New York has many small resorts. Northern states generally offer more night skiing to make up for shorter daylight hours. However, these state features alone did not clearly predict prices. Principal Component Analysis (PCA) showed that states with similar ski markets do not cluster by ticket price.

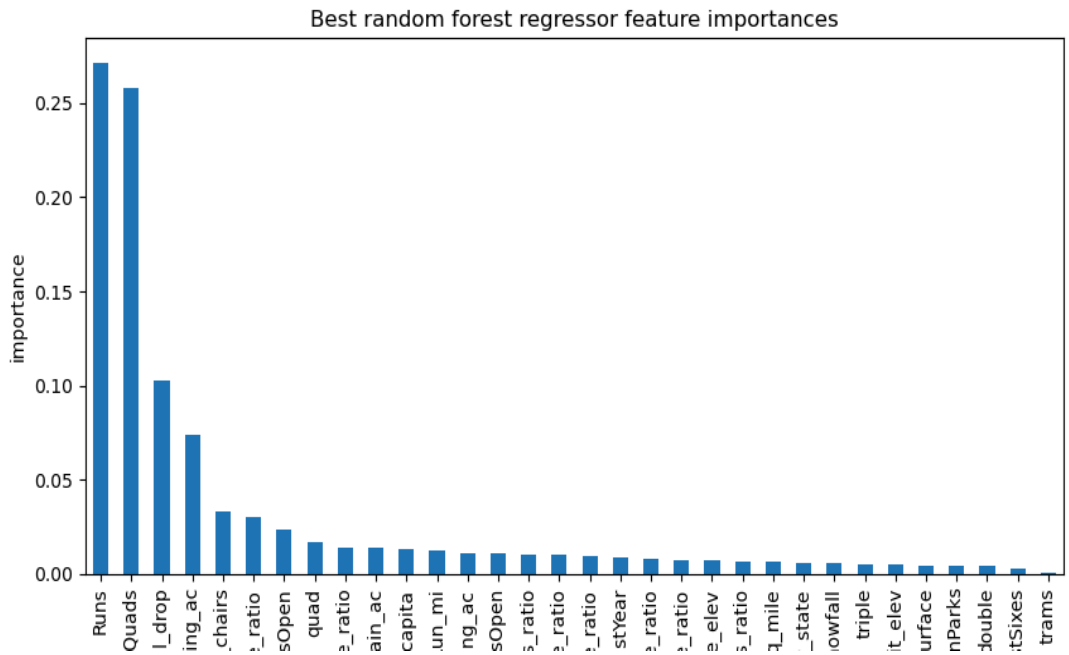
Modelling:

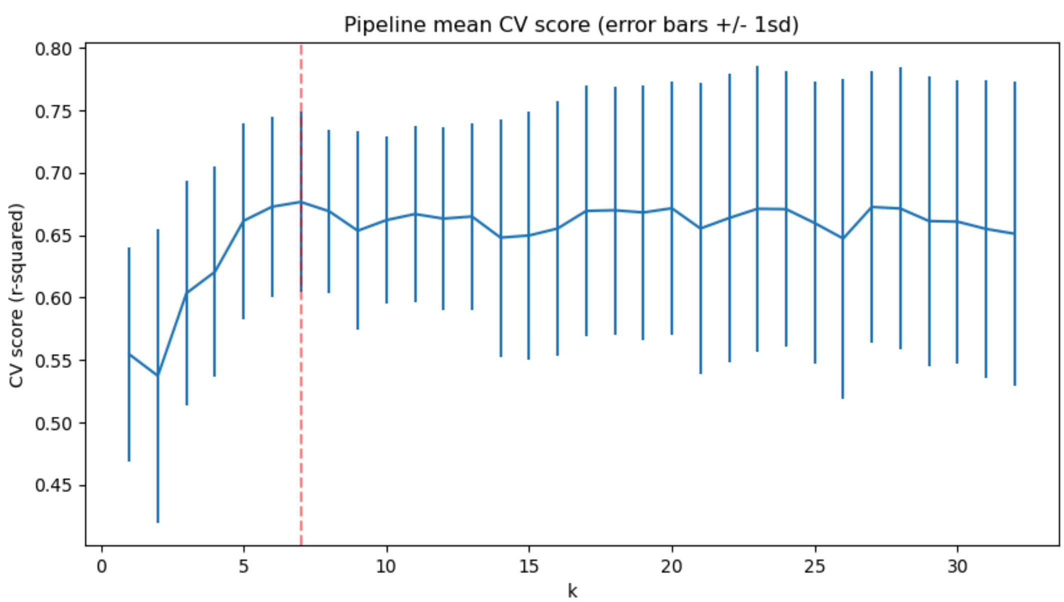
We began with a baseline model that predicted the average adult weekend ticket price for all data points. As expected, this simple model yielded an R^2 of zero on the training set and slightly negative on the test set, indicating no real predictive power beyond guessing the mean. The mean absolute error was about $19, setting a basic benchmark for performance.

Next, we developed an initial linear regression model using all numeric features. To prepare the data, missing values were imputed with the median, and features were standardized to zero mean and unit variance. This model showed a strong improvement with an R^2 around 0.83 on training data and 0.71 on the test set, along with an average error near $9. However, the gap between training and testing performance suggested some overfitting, highlighting the need for feature refinement.

To address this, we applied feature selection with SelectKBest, tuning the number of features using GridSearchCV. The optimal model used about eight features, including key variables like vertical drop, snow-making coverage, and lift capacity, all aligning with domain expectations. This refinement reduced model complexity while maintaining predictive power. The cross-validated R² averaged 0.75 with low variability, and test results confirmed the model's robustness.

<u>Below are the figures showing Feature Importance and SelectKBest scores:</u>





To capture potential nonlinear relationships and interactions missed by linear models, we trained a Random Forest Regressor. We found median imputation still worked best, and scaling didn't significantly affect performance. After tuning hyperparameters via GridSearchCV, the Random Forest slightly outperformed the refined linear model, reducing the mean absolute error by about $1 and showing more stable performance across folds. Test results were consistent, reinforcing confidence in its predictions.

Scenario Modelling :

The Random Forest Regressor gave the best results, with the lowest RMSE and the highest R^2. Using this model, we ran different scenarios to see how changes might affect the ticket price. The results showed:

- Closing up to 3 runs has almost no effect on ticket price, but closing more than 4 or 5 runs starts to lower the price significantly.
- Adding more vertical drop and a new lift increases the ticket price by only about $0.41 per ticket, which does not cover the high cost of building the lift.
- Adding extra snowmaking or extending runs also has little or no impact on the supported price.

Based on the model, Big Mountain Resort appears to be underpricing it's tickets. Given its strong facilities and how similar resorts price their tickets, the analysis suggests that the resort could charge closer to $92–$97, compared to the current $81. We recommend increasing the ticket price gradually instead of making a big jump all at once. It would help to show clear comparisons to other similar resorts to explain why the price increase is fair. Any price increase should also be paired with visible guest experience improvements and good marketing to keep customers happy and loyal.

In conclusion, the project shows that Big Mountain's strong existing facilities do support a higher ticket price, but large expansions like new lifts do not increase price enough to pay for themselves through ticket sales alone. Therefore, the resort should focus on getting the most value out of what it already has and consider smaller, cost-effective upgrades that improve the guest experience.