Project Summary

# Problem statement:

Can Big Mountain Reports be able to recoup a $1.5 million increase in operating costs by raising the average ticket price by $5 by the end of the upcoming ski season or by April 15th, 2026, considering the features and facilities it offers for skiers and riders of all levels and abilities?

Data Wrangling and EDA:

The raw CSV file was read and converted into a dataset containing 330 resorts with 27 features. Initial information about the dataset was obtained using various methods. Missing values, duplicates, and data distribution were explored. Extreme values were corrected, and a new dataframe was created using the group by function and aggregate methods. The weekday ticket price column was dropped due to a higher number of null values, resulting in a higher overall ticket price.

Feature Engineering:

The cleaned ski data was joined with statewide summary data to enhance feature comparison and correlation. Feature engineering was performed to calculate the ratio of each resort within its state. Principal Components Analysis (PCA) was applied after scaling the data to reduce the number of features while retaining variance. The combined dataset was evaluated for correlations using a seaborn heatmap, identifying AdultWeekend ticket price correlations with features like fastQuads, Runs, Snow Making\_ac, resort\_night\_skiing\_state\_ratio, and total\_chairs. Scatter plots were used to visualize these correlations, highlighting prominent features for training and modeling.

Model Preprocessing:

In this stage several operations were performed to create the best model. Big Mountain data was excluded to allow the model to learn from other data, reserving 30% for testing. The 'Adultweekend' column was used as the dependent variable, while the remaining columns served as independent variables. Metrics such as R squared, Mean Absolute Error (MAE), and Mean Squared Error (MSE) were calculated. The first model, sklearn's Dummy Regressor, served as a placebo control group. Standard Scaler was applied to standardize the dataset. Linear Regression on scaled data improved the R squared to 0.720 and reduced the MAE to 9.407. Pipelines in SKlearn with Simple Imputer, Standard Scaler, and Linear Regression yielded similar results. SelectKBest was used for feature reduction, resulting in R squared = 0.625 and MAE = 11.201. Cross-validation optimized model parameters, producing a mean R squared score of 0.632. Hyperparameter tuning with GridSearchCV identified the best values for K in SelectKBest. Random Forest Regressor was applied, and the final model with Random Forest Regressor achieved an MAE of 9.537. The random forest model demonstrated lower cross-validation mean absolute error and less variability.

The final model was selected with Random Forest Regression using optimal parameters. It is then saved as a pipeline file for future use and application.

# Recommendation:

After developing an optimal model to predict ticket prices, and calculating for Big Mountain Resorts, the model's findings and recommendations are:

The current price for Big Mountain Resorts is **$81** per ticket1. The model's recommended price is **$95.87**. Considering the Mean Absolute Error of **$10.39**, the suggested new price range would be **$85.87** to **$105.87**.

|  |  |  |  |
| --- | --- | --- | --- |
| Current Price | Recommended Price | Mean Absolute Error | Implication |
| $81 | $95.87 | $10.39 | Increase in ticket prices |

# Comparison of Big Mountain with Other Resorts:

Big Mountain offers one of the highest vertical drops, extensive snow making, numerous chairs and runs, the longest run, and substantial skiable terrain.

|  |  |
| --- | --- |
| Category | Percentile Score Nationwide |
| fast Quads | 93.50% |
| quad chairs | 81.77% |
| triple chairs | 98.01% |
| total chairs | 94.04% |
| Number of Runs | 93.07% |
| Terrain Parks | 82.83% |
| Longest Run | 95.96% |
| Skiable Terrain | 98.36% |

A graph of a vertical drop

AI-generated content may be incorrect.A graph of a graph

AI-generated content may be incorrect.

A graph of chairs distribution

AI-generated content may be incorrect.A graph with numbers and lines

AI-generated content may be incorrect.

A graph with a red line

AI-generated content may be incorrect. A graph of a number of areas

AI-generated content may be incorrect.

-   While Big Mountain has the **highest rates in Montana**, there is potential for further price increases given the superior facilities it offers compared to other establishments in the state.

A graph with blue lines

AI-generated content may be incorrect. A graph of a number of blue bars

AI-generated content may be incorrect.

Scenario Analysis for Big Mountain

# Potential Improvements and Their Impact

## Scenario 2

* Modifications: 1 additional run, increase vertical drop by 150 feet, install 1 chair lift
* Expected Increase in Ticket Price: $8.61
* Expected Revenue Increase: $15,065,471

## Scenario 3 and 4

* Scenario 3: Add 2 acres of snowmaking
* Scenario 4: Extend the longest run by 0.2 miles and add 4 acres of snowmaking capability
* No notable changes in ticket prices or revenue

## Scenario 1

* Risk: Closing more than 5 runs could significantly reduce ticket prices and revenue for future seasons
* Recommendation: Test the impact of closing up to 5 runs to minimize risk

# Assumptions:

* This recommendation is based on the data provided to our data analysis department and assumes that the data is valid and current.
* It also assumes that the recently installed lift is included in this dataset.
* Additionally, it is assumed that all features affecting ticket prices are included in this dataset and considered by other resorts when estimating their ticket prices.
* We expect the upcoming season to attract 350,000 visitors, with an average stay of 5 days.

# Data Recommendation:

To enhance future evaluations and improve modeling, we recommend upgrading our data collection capabilities, acquiring more comprehensive market business data, and investing in the latest data tools and hardware to make more informed decisions.