Big Mountain Resort Ticket Pricing Analysis

Executive Summary

Big Mountain Resort currently charges $81 for an adult weekend ticket. Our random forest model, trained on peer resort data, predicts an optimal price of $100.55, revealing a potential undervaluation. This $19.55 gap suggests an estimated missed revenue of $34.2M given 350,000 annual visitors averaging 5 tickets each. The analysis highlights infrastructure features such as vertical drop, lifts, and runs as the strongest drivers of ticket prices. Scenario testing indicates that adding runs, vertical drop, and lifts significantly improves revenue potential, while snowmaking expansions show minimal effect.

Data Wrangling & Cleaning

The raw dataset contained approximately 330 U.S. ski resorts with 27 features. Initial steps included auditing for missing values, duplicates, and inconsistent data types. Key issues identified were:

- Missing values in ticket price features (AdultWeekday and AdultWeekend) and lift-related fields.
- The 'fastEight' feature contained over 50% missing values and was dropped.
- Outliers were detected and corrected, such as Silverton Mountain's skiable terrain (26,819 acres corrected to 1,819).
- Implausible values like '2019 years open' were corrected.

Categorical checks revealed duplicate resort names (e.g., Crystal Mountain), resolved by combining with state identifiers. Region vs. state variables were validated, showing region adds useful clustering information beyond state boundaries. Finally, external state-level population and area data were merged to enrich the dataset with demand-side context.

Exploratory Data Analysis (EDA)

EDA highlighted key distributional and relational insights:

- Ticket prices varied widely, with most resorts between $40–$100, but some above $150.
- States like Colorado, Utah, and Vermont had higher average prices, while Midwestern states were lower.
- Vertical drop, snowmaking acreage, number of runs, and fastQuads correlated strongly with higher ticket prices.
- PCA showed Vermont and New Hampshire stand out due to unusually high resort density.

These findings suggested that resort-level features are stronger predictors of ticket price than state labels.

Preprocessing

The dataset was split into training (70%) and testing (30%) sets. The target variable was AdultWeekend ticket price, leaving 33 numeric predictors. Preprocessing included:
- Median and mean imputation tested for missing values, with median performing slightly better.
- Scaling via StandardScaler ensured comparability across features.
- Pipelines were constructed to streamline imputation, scaling, feature selection, and modeling steps.
Cross-validation (5-fold) was used to estimate model generalization, preventing overfitting.

Modeling Methodology

We built and compared several models:
- Baseline: average ticket price across resorts (poor accuracy).
- Linear Regression: feature selection via SelectKBest.
- Random Forest Regressor: chosen due to higher accuracy, lower mean absolute error, and stability across folds.
Preprocessing included median imputation, optional scaling, and grid search over key hyperparameters.

Results

Model	Mean Absolute Error (Test)	Key Notes
Baseline (Average)	~$15+	Benchmark only, poor performance
Linear Regression	≈ $11.8	Selected features improved interpretability
Random Forest	≈ $10.0	Best performing model, more stable

Feature importance (RF): Vertical drop, fastQuads, runs, and snowmaking area were dominant drivers.

Scenario Analysis

Scenario	Description	Impact
1	Close up to 10 runs	Negligible until >2 closures; large revenue drop beyond 5

2	Add 1 run, 150ft vertical, +1 chair	+$2.25 per ticket, ≈ +$393,750 revenue
3	Scenario 2 + 2 acres snowmaking	No significant additional effect
4	Add 0.2 miles run + 4 acres snowmaking	No measurable effect

Recommendations

- Adjust pricing upward to align closer with $100.55 model-predicted value.
- Prioritize investments in vertical drop, runs, and lifts for measurable revenue impact.
- Deprioritize snowmaking expansion as it shows limited contribution to price.
- Use the random forest model iteratively for strategic decision-making.

Limitations & Further Work

The analysis relies solely on facility-based features and assumes peer resorts set prices fairly. Operational costs, demand elasticity, and brand effects were not included. Future work should integrate cost data, customer willingness-to-pay surveys, and pricing tier structures for more holistic recommendations.