**Big Mountain Resort Pricing Analysis Report**

**Problem Statement:**  
Increase Big Mountain Resort’s total revenue by 5% within the next ski season by optimizing ticket pricing based on data-driven insights and considering guest preferences, while improving guest satisfaction scores.  
  
**Context:**Big Mountain Resort, a popular ski resort in Montana, is looking to increase its revenue. They believe their current pricing strategy, based solely on market averages, is not fully capitalizing on their unique offerings. The resort wants to implement a more data-driven approach to pricing to better understand guest preferences and optimize revenue.

**Data Wrangling:**  
In the data wrangling section, we systematically processed and prepared the ski resort data to ensure its readiness for analysis. Below are the key steps and findings:

* Data Exploration and Initial Inspection  
  - We began by loading and exploring the ski resort dataset, focusing on identifying our resort of interest and understanding the overall structure of the data.

- A thorough review of the dataset revealed the presence of missing values across various columns, particularly in the ticket price data, which is crucial for our analysis.

* **Categorical Features Analysis**  
  - We examined categorical features, including resort names, regions, and states.  
   - Through this, we identified duplicated resort names and confirmed that combining these with state or region information provided unique identifiers.

A graph of different colored columns

Description automatically generated with medium confidence

- We also analyzed the distribution of resorts and ticket prices across different regions and states, which helped us understand regional trends and pricing patterns.

* **Numeric Features Analysis**  
  - The numeric features were summarized, and their distributions analyzed.  
   - Specific focus was placed on features like `SkiableTerrain\_ac`, `Snow Making\_ac`, and lift types (`fastEight`, `fastSixes`, and `Trams`).

A screenshot of a graph

Description automatically generated  
- This analysis revealed inconsistencies, such as unusual data entries for certain resorts, which were addressed accordingly.

* **State-wide Summary Statistics**  
  - We derived state-wide summary statistics to better understand the market segment.  
   - This included aggregating relevant numeric features at the state level to provide a broader context for each state’s ski industry.
* **Data Cleaning**  
  - Rows with no ticket price data were dropped, ensuring that our dataset was free of critical missing information that could impact the analysis.  
  - Distribution reviews were conducted post-cleaning to verify that the data was consistent and representative.
* **Population Data and Target Feature**  
  - Population data was incorporated to provide additional context for the target feature, which is the ticket price.  
   - This helped refine our understanding of the relationship between state demographics and resort pricing.  
  - We also assessed the missing values by resort, ensuring that any resorts with significant gaps in data were either corrected or excluded.
* Adult Weekend vs Adult Weekday depicts no difference between them to suggest that Target feature to be Adult Weekend   
  A graph of blue dots

  Description automatically generated

**Exploratory Data Analysis:**  
In this exploratory data analysis, we undertook a comprehensive examination of both the ski resort data and the state-wide summary data. Our analysis covered several key areas, which provided insights into the factors influencing ticket prices and resort characteristics across various states.

* **Top States by Summary Statistics**  
  - We identified the leading states by various metrics such as total area, population, number of resorts, skiable area, night skiing area, and days open.  
  - This analysis highlighted states like New York, California, and Colorado as dominant in these categories.
* **Resort Density**  
  - By calculating resort density, we identified states with the highest concentration of ski resorts relative to their geographic area.  
  - This provided a clearer understanding of the regional distribution of ski resorts. Some states are higher in some but not in others
* **Data Scaling and PCA**  
  - We scaled the data and applied Principal Component Analysis (PCA) to reduce the dimensionality and visualize the data more effectively.  
  - This heatmap shows correlations between target feature, Adult Weekend with fastQuads, Runs, SnowMaking\_Ac and resort\_night\_skiing\_state\_ratio

A close-up of a red and pink pixelated pattern

Description automatically generated

**Model Preprocessing and Feature Engineering:**  
In the model preprocessing phase, we ensured that the dataset was prepared for modeling through several key steps

- Feature Selection: Based on the exploratory analysis, we selected features that showed significant correlations with ticket prices.  
- Data Splitting: The dataset was divided into training and testing subsets to facilitate model evaluation.  
- Linear Regression:

* + This algorithm was used to establish baseline predictions based on the selected features.
  + Regression analysis with mean and median as strategy were performed.
  + Mean Absolute error was off by around $19 using mean, while with median error was only off by $9.
  + During the analysis, the vertical drop is found to be the biggest positive feature.

- Random Forests:

* An ensemble method that improved prediction accuracy by aggregating results from multiple decision trees.
* Random forest model has a lower cross-validation mean absolute error by $1 with less variability

A graph with blue and white text

Description automatically generated

**Modelling:**

The model predicts a ticket price of $95.87 for Big Mountain Resort, compared to the current price of $81.00. Even accounting for a mean absolute error of $10.39, this indicates potential room for a price increase. To determine a fair price increase we looked Big Mountain Resort (represented by dashed red line) top features in market context.

* Vertical DropA graph of a vertical drop

  Description automatically generated
* Snow making area A graph of snow making

  Description automatically generated
* Total number of chairs

A graph with blue and red lines

Description automatically generated

* Fast quads

A graph with numbers and lines

Description automatically generated

* Runs

A graph of a number of runs

Description automatically generated

* Longest run

A graph of a running graph

Description automatically generated

* Skiable terrain area

A graph of a number of land

Description automatically generated

Big Mountain Resort either ranks high or well above the average in each category, this tells us that a big mountain resort is great resort with lot of great facilities and the price should reflect that.

**Modelling Scenarios:**

Big Mountain Resort has been reviewing potential scenarios for either cutting costs or increasing revenue (from ticket prices). Ticket price is not determined by any set of parameters; the resort is free to set whatever price it likes. However, the resort operates within a market where people pay more for certain facilities, and less for others. Being able to sense how facilities support a given ticket price is valuable business intelligence. This is where the utility of our model comes in.

The business has shortlisted some options:

1. Permanently closing up to 10 of the least used runs. This doesn't impact on 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 cover

4. Increase the longest run by 0.2 mile to boast 3.5 miles length, requiring additional snow making coverage of 4 acres

The expected number of visitors over the season is 350,000 and, on average, visitors ski for five days. Assume the provided data includes the additional lift that Big Mountain recently installed.

In evaluating the potential scenarios for Big Mountain Resort, the model provides key insights:

1. **Closing Runs**: Closing one run has no impact on ticket prices, but closing 2-3 runs reduces support for pricing. Closing 4-5 runs results in no further losses, while closing 6 or more leads to a significant drop-in ticket price support.

A graph of a price

Description automatically generated with medium confidence

1. **Adding a Lower Run (Scenario 2)**: This option increases ticket price support by $1.99, potentially generating an additional $3.47 million in revenue over the season.
2. **Adding Snowmaking Coverage (Scenario 3)**: Like Scenario 2, this increases ticket price support by $1.99, with the same projected revenue boost of $3.47 million.
3. **Extending the Longest Run (Scenario 4)**: This has no impact on ticket price, as the longest run is not a significant factor in the final model.

Overall, adding a lower run with or without additional snowmaking offers the most significant potential for revenue growth, while closing runs should be approached cautiously due to the negative impact on pricing.