# Guided Capstone Project Report

## Problem Statement

What action plan of increasing ticket prices, cutting costs, or capital investments can Big Mountain Resort take to boost revenue sufficient to cover the $1.54 million increase in operating cost and achieve an 8 percent profit by the end of this season.

## Data Wrangling

The database team has provided a .csv file with tabular data on 330 ski resorts across the country. It contains 27 columns with features describing each resort including geographic location, ticket price, amenities and various characteristics of the facilities. For the sake of data quality and relevance, we performed careful modifications to the data, including the following …

- “fasteight” - column removed - too many NULLs
- “weekdayprices” - column removed - incomplete data. “weekendprices” data was more complete and used instead
- rows removed - any resort without price information was removed as that is the key feature
- …and more

## Exploratory Data Analysis

We reviewed the top states for key summary statistics including state size, population, # of resorts per state, total skiable area, total night-skiable area, total days open, resort density (# of resorts per 100k capita). Also used PCA to simplify the dimensionality of the dataset’s many variables to a more manageable size. There are a handful of reasonably strong correlations between ticket price and runs, fast_quads, snow_making_ac, etc. A derived feature "resorts_per_100ksq_mile" was found to be an especially good target feature for correlating ticket price. Our analysis determined it would be best to treat states equally.

## Model Pre-processing with Feature Engineering

We took our imported data and split it into isolated train and test partitions. Train set would be used for *training* the model to help it fit the ideal values for our model and the test set would be used to *test* if our model generalizes well on data it was not trained on. We took the mean to use as a benchmark for whether our model met a minimum performance standard. We further evaluated models with criteria such as R squared, Mean Absolute Error, and Mean Squared Error. We created pipelines to group our processing steps together and check our models (Linear regression and Random Forest) to see which optimized performance by minimizing the

error as this would mean most accurate results. Missing values were imputed using the data set median to ensure the model wasn't led astray by ambiguous or outlier data.

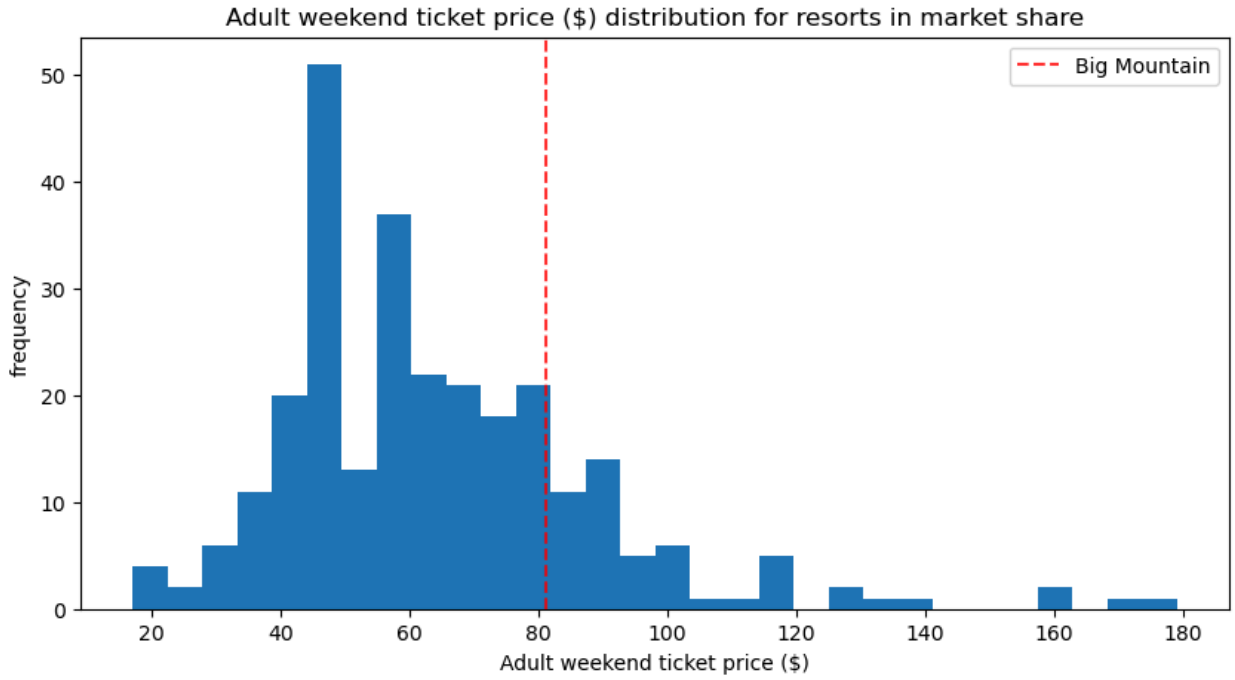
## Model Selection

A Random forest model was chosen as it was more accurate than the alternate Linear Regression model by about a dollar, MAE, and had a lower variance.

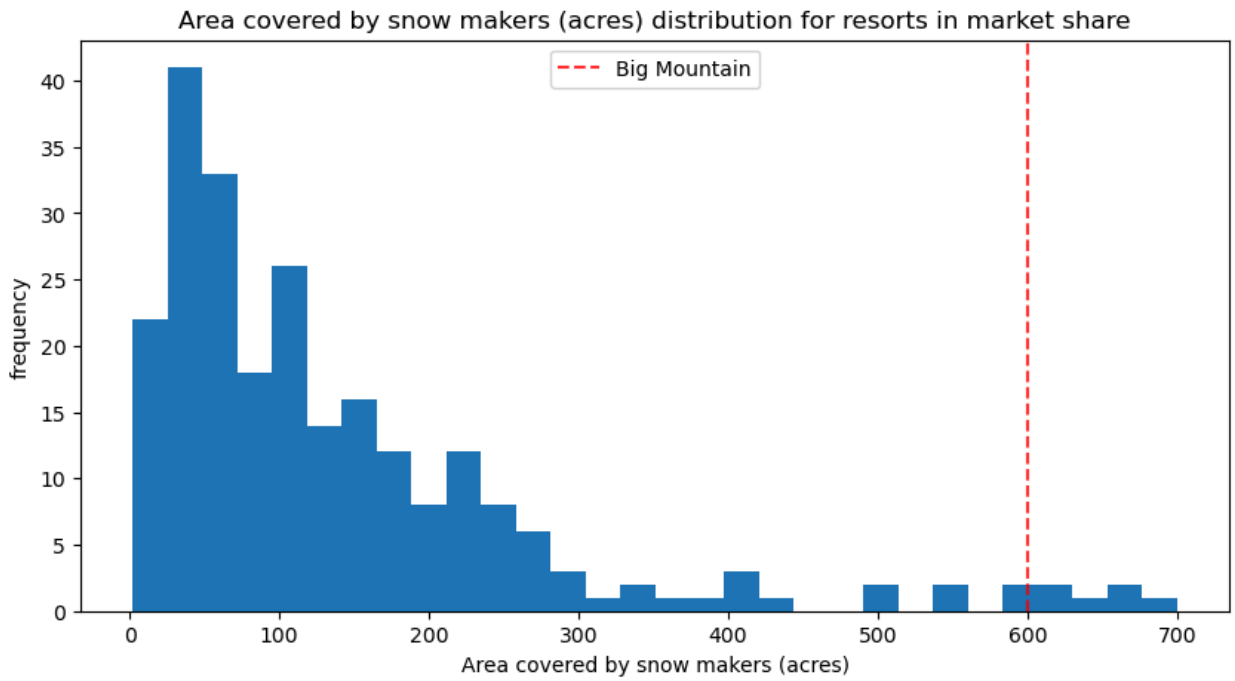
## Pricing Recommendation and Conclusion

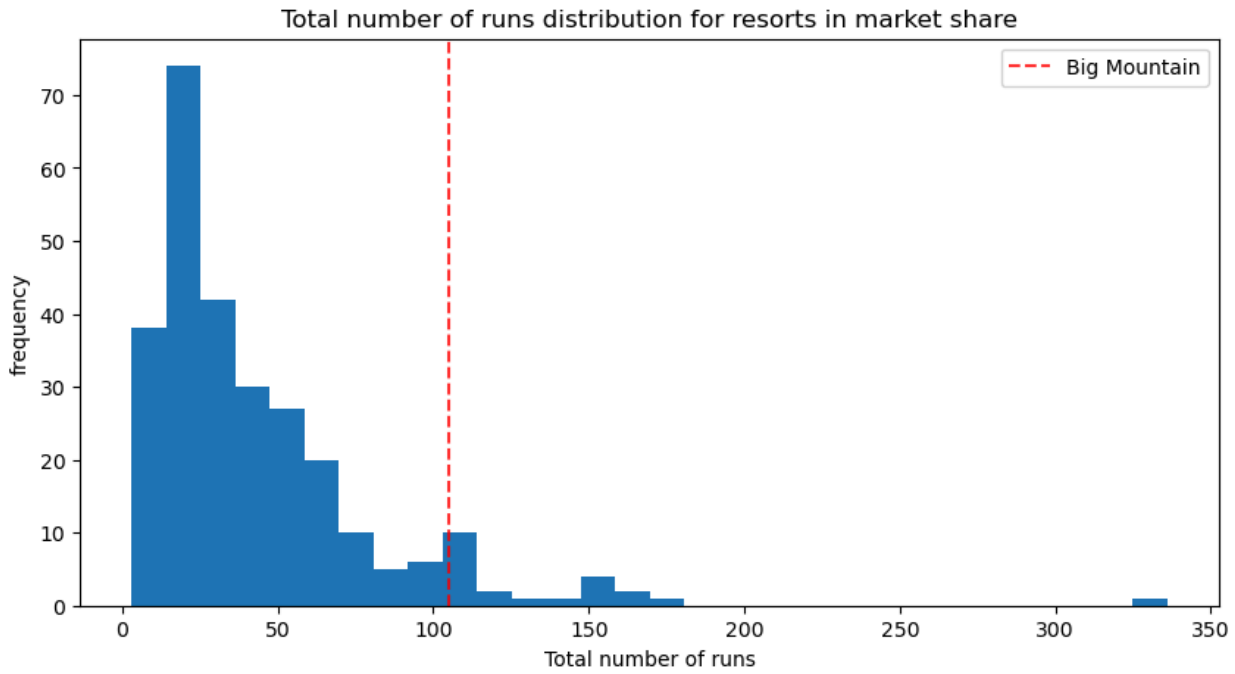
The data and analysis supports increasing the ticket price due to Big Mountain's status as a market leader based on their facilities. The new recommended ticket price is $95.87.

Currently, Big Mountain sits roughly in the upper mid portion of ski resorts with its ticket price.

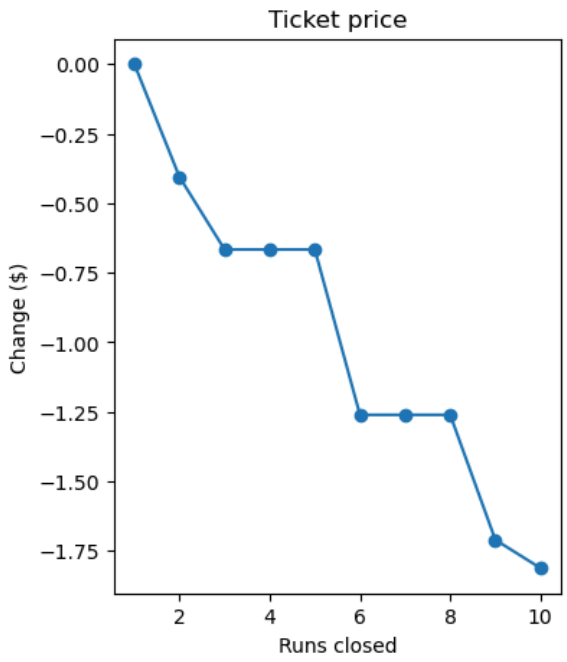


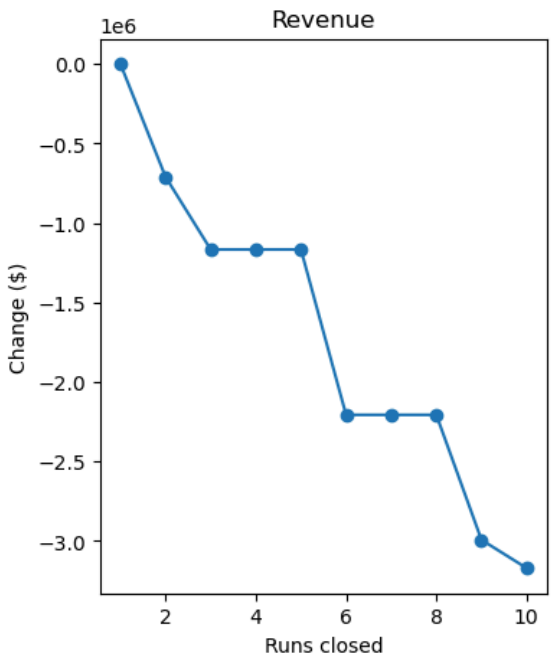

However, for some of the most strongly correlated features with higher ticket prices (e.x. snow making area and total # of runs), Big Mountain (represented by the red line) stands near the lead of the pack….




In addition to increasing revenue via ticket price increase, the analysis shows that one run can be closed, saving on operating costs but without any consequence to resort revenue.




## Further Scope of Work

Going forward, it would be useful to have more information to evaluate ticket revenue modeling scenarios with. Example data that would be useful…

- # of ticket sold and when (perhaps tickets sold per month of each year to track customer demand)
- complete ticket prices for both adult *and* children for weekend *and* weekday
- further granularity breakdown of operating costs (utilities, labor, snow equipment maintenance, etc)

That said, this is a nice-to-have, and it would not be our recommendation to pursue if the costs of collecting and storing the information was high. The existing data is sufficient for revenue modeling. We'd recommend taking the price increase suggestion and comparing how sales and profit benefits over the next season accordingly.

The insights from this work can be further leveraged using our SAS-hosted modeling tool. The tool provides a hands-on dashboard UI allowing feature selection and parameter value input which Big Mountain or their consultants can use to perform custom ad-hoc analysis to guide and support business revenue modeling decisions.