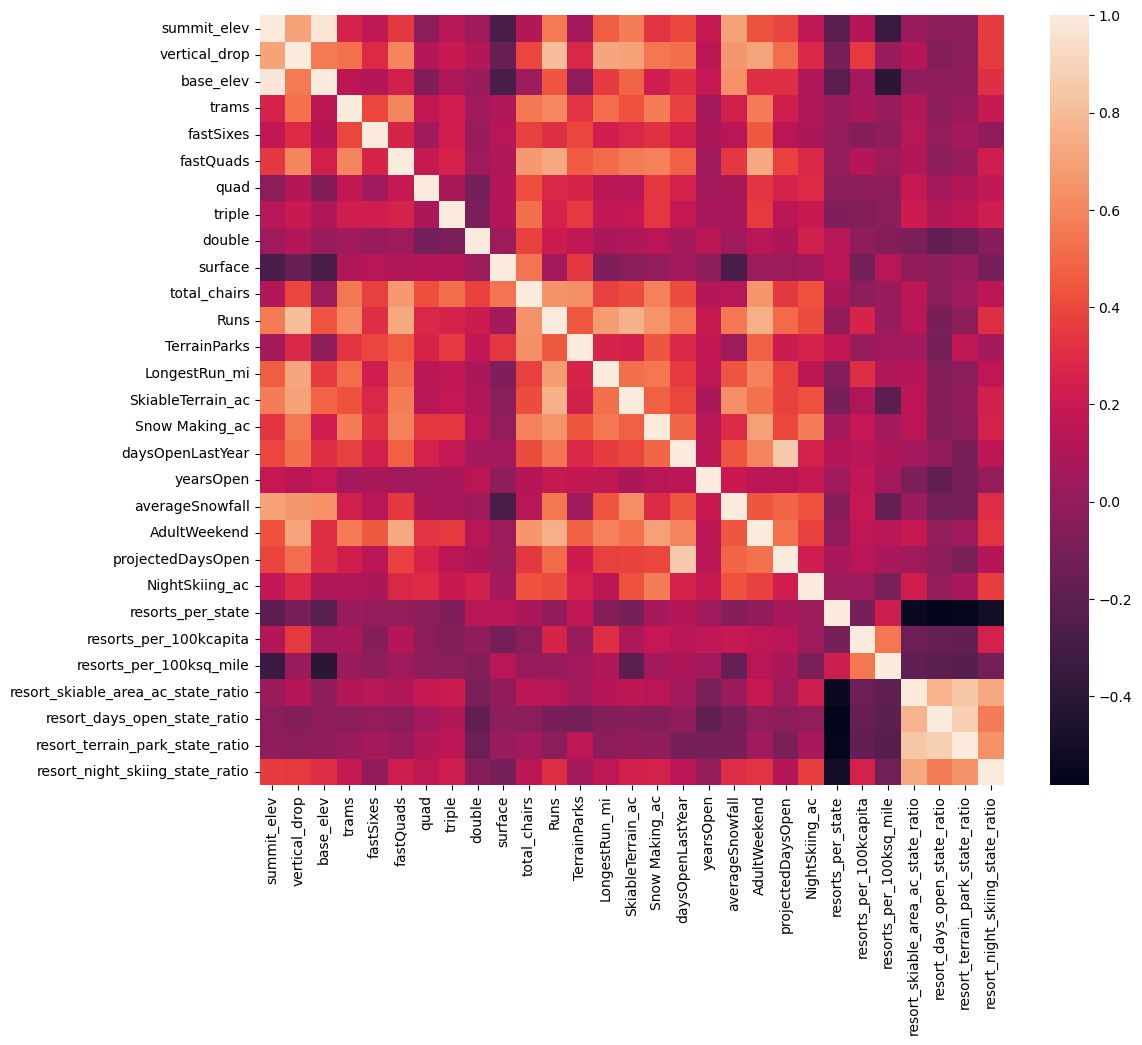
Guided Capstone Project Report

The initial question in mind consisted of the problem statement: “If Big Mountain Resort implements a dynamic, data-driven pricing strategy based on demand patterns, competitor pricing, and historical data, will it increase overall revenue by at least 15% without reducing customer demand by the end of December, 2025?”. A little context is the Big Mountain Resort has recently added a new chair lift, increasing operational costs by $1,540,000 for the season. The resort wants to implement a data-driven approach to ticket pricing and operational efficiency.With this in mind, I worked on creating a model of what facilities matter most to visitors, particularly which ones they're most likely to pay more for, in order to start the modeling we needed to make sure we were working with only necessary data. I cleaned and prepared the ski resort dataset for predictive modeling. The original file (ski\_resort\_data.csv) had 330 entries, each representing a ski resort in North America—including our resort of interest, Big Mountain Resort. After cleaning, the final dataset (ski\_data\_cleaned.csv) has 277 rows—a 16% reduction from the original. This version preserves the important information while being more suitable for building a predictive model.

With the data preparation/wrangling complete, I was ready to move on to exploratory analysis and feature selection. First I began searching for the features of the data that showed correlation using PCA(Principal component analysis), I was able to scale the data according to the variance and find the correlations that truly mattered for the adult weekend ticket price, my go to graph for this would be the heatmap, however I created scatterplots of the ticket prices compared to each numeric variable in order to help prove the correlations directly to ticket prices.



From this graph you can see that the adult weekend prices are closely correlated with fastQuads (ATVS), snow Making AC, Runs, Total chairs. This shows that the more snow and the more they are able to ski the more likely they will be willing to pay more leading to more revenue. These correlations led to the epiphany that people seem to put more value in guaranteed snow cover rather than more variable terrain area. In the end I chose the average weekend ticket price as the target to predict for our model but I began with preprocessing and training the model.

We began by establishing a baseline using the average weekend ticket price, which performed poorly but provided a reference point. A linear regression model was then built with preprocessing steps like imputation, scaling, and feature selection. It showed moderate accuracy in cross-validation, and its test performance was consistent, identifying features like vertical drop and snowmaking as important predictors. We then trained a random forest regressor with tuned hyperparameters. It required minimal preprocessing and outperformed the linear model both in cross-validation and on the test set. Feature importance revealed vertical drop, number of runs, and projected days open as key drivers. Based on its higher accuracy and reliable performance, we chose the random forest model for the next phase to support data-driven business decisions.

After modeling I believe scenario 2 is the best bet, to 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.

Big Mountain currently charges about $110 per day ticket. Modeling suggests a ticket price of $120-$130 is feasible due to new facilities, with a $3.00 increase in support for the ticket price from this scenario, potentially yielding $5,250,000 over the season. Proposing this to leadership requires highlighting the enhanced value and providing data-driven justification. The new chairlift's operating cost adds only $0.50 per ticket, easily absorbed by the price increase. Notably, a small increase in snowmaking area makes no significant difference to the supported ticket price in this model. For future improvements, prioritize snowmaking expansion (beyond minor additions) and optimizing the longest run. Any run closures should be piloted temporarily after data analysis, with clear guest communication and careful monitoring of impact. Big Mountain's modeled price is significantly higher than its current price, a likely surprise to executives, indicating potential underpricing given its strong facilities.

|  | **vertical\_drop** | **Snow Making\_ac** | **total\_chairs** | **fastQuads** | **Runs** | **LongestRun\_mi** | **trams** | **SkiableTerrain\_ac** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **124** | 2353 | 600.0 | 14 | 3 | 105.0 | 3.3 | 0 | 3000.0 |