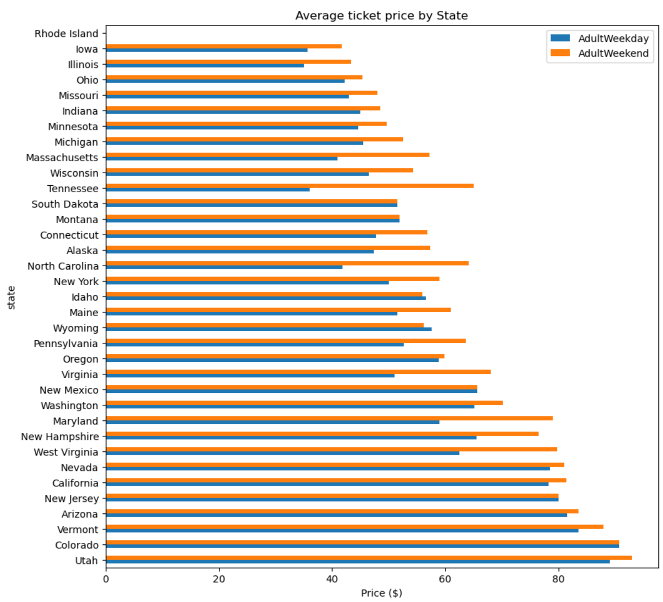
GUIDED CAPSTONE - PROJECT REPORT

The intent of this analysis was to find a way to allow Big Mountain Resort to neutralize a $1,540,000 increase in operating costs this season. I hypothesized that, by better capitalizing on its facilities, the resort could create a better value for the current ticket price, or even enable a higher ticket price.

The initial dataset contained 330 rows, including our resort. Notably, skiable area data for Silverton Mountain was corrected due to extreme outliers. The 'fastEight' column with minimal information was removed, along with a row with erroneous 'yearsOpen' data. Rows with missing price data (around 14%) were dropped, leaving 277 rows.

The target feature for predicting ticket price is 'AdultWeekend' – the weekend ticket price – chosen for its availability of data. Interestingly, both weekend and weekday prices were the same for all Montana resorts.



A key insight was the price variation across states. To simplify analysis, resort data was aggregated by state, as similar states tend to have comparable prices. By integrating Wikipedia data on state-level population and area, additional relevant features were added for future price predictions.

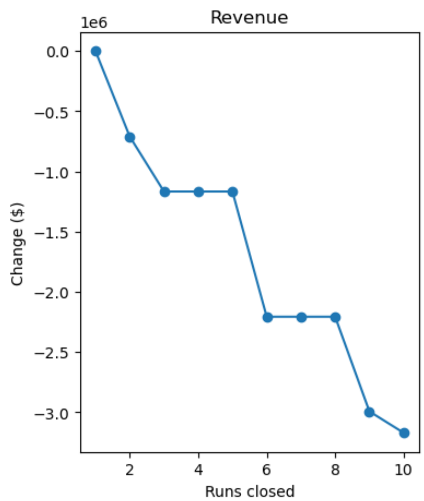
During the exploratory data analysis phase, I delved into both state-level and resort-level information to inform my analysis strategy. Notably, I addressed the treatment of state labels and observed that the relevant features for price analysis remained consistent across states after accounting for population and area differences through PCA. This led me to prioritize numerical over categorical factors for subsequent modeling, with AdultWeekend chosen as a suitable proxy for price.

The analysis highlighted key resort-level features that exhibited positive correlation with price, including vertical drop, fastQuads, Runs, Snow Making_ac, resort night skiing state ratio, resorts per 100k capita, and total chairs. To enhance the analysis's clarity and mitigate multicollinearity, dimensionality reduction was performed by eliminating redundant features. However, potential challenges emerged, such as high correlations between certain constructed features and their constituent elements, necessitating careful consideration of multicollinearity in the modeling phase.

Many crucial insights emerged during preprocessing and training. Initially, the baseline approach using average price yielded an MAE of about 19, which serves as a reference point for subsequent models.

A linear model was developed, revealing that highly predictive features mirrored those strongly correlated with price in the previous analysis. Vertical drop, Snow Making_ac, total_chairs, fastQuads, Runs, and LongestRun_mi positively impacted price, while trams and SkiableTerrain_ac had negative associations. The cross-validated MAE for the linear regression model was approximately 11.8, consistent with the test split performance (around 10.5).

A random forest model was also constructed for comparison. Four influential features for price prediction were identified: fastQuads, Runs, Snow Making_ac, and vertical_drop. Cross-validation estimated a superior MAE of about 9.5 for the random forest model compared to linear regression. Ultimately, the decision was made to proceed with the random forest model despite its potentially more complex interpretation. The notable performance improvement justified this choice, enhancing our predictive capabilities for guiding critical business decisions.



The modeling outcomes yielded some valuable insights for strategic decision-making. Big Mountain's current ticket price is $81.00, while my modeling suggests a sustainable price of $95.87 considering the marketplace. This price adjustment could be proposed to leadership, supported by significant resort comparisons. In the case of the new chair lift scenario, while there are operating costs of approximately $1.5 million, the model predicts a revenue increase of nearly $3.5 million over the season. By raising the ticket price by $2, not only can the new expenses be covered, but an additional $2 million profit can also be generated.

My analysis reveals that the original hypothesis, that the resort can sustainably increase its ticket price by a substantial margin, was well-founded. I concluded that Big Mountain’s heightened operating costs could be mitigated by better capitalization of existing facilities. In terms of future scope, the model I created can be a powerful tool for ongoing decision-making. Rather than consulting directly for each parameter combination, the model can be adapted to facilitate versatile exploration by business analysts. Additionally, by incorporating various variables such as annual visitors, length of stay, and operating costs, the model's predictability can be extended, enhancing its utility for strategic analysis.