Guided Capstone Project Report

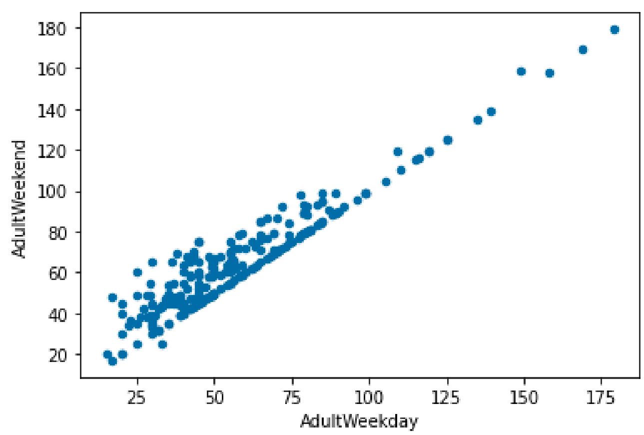
Overview

An analysis of market share ski resorts was conducted to determine an appropriate weekend ticket price for Big Mountain Ski Resort in Montana. **The analysis revealed that a price increase could be warranted - from $81.00 to $95.87 per ticket - with an expected increase in annual revenue of $26 million**. This report provides an overview of the methodology used to develop the recommendation along with additional potential cost-saving and revenue-increase scenarios.

Methodology

Data Wrangling

One of the major goals of wrangling the data was to determine whether the weekend or weekday ticket price should be used in the model. After review the cleaning the original data set and merging additional data for population and state data, summary statistics and scatter plots were developed to understand the relationship between weekend and weekday ticket prices:



Ultimately, the weekend price was chosen to be used in the model. The similarity in weekend vs. weekday prices for resorts above $100, the similarity of prices in Montana, and the greater number of data points for weekend tickets was compelling enough to choose weekend ticket prices over weekday ticket prices.

Exploratory Data Analysis

Exploratory data analysis was conducted to look for patterns between states and determine if the state label would be useful in the modeling stage of the overall analysis. A variety of explorations were undertaken, including reviewing summary statistics for several resort features, identifying resort density, and principle components analysis to decrease the number of features. A total of eight new features were developed (four for resort competition and four for ease of transportation) and several features were seen to have a correlation with ticket price, including the number of fast quads, runs, snow making ability and total chairs (see [Appendix A](#) for visuals, including a feature heat map and scatter plots comparing ticket price with features). **Ultimately, it was found that no clear patterns emerged between states and that the state label was not particularly useful; all states should be treated the same.**

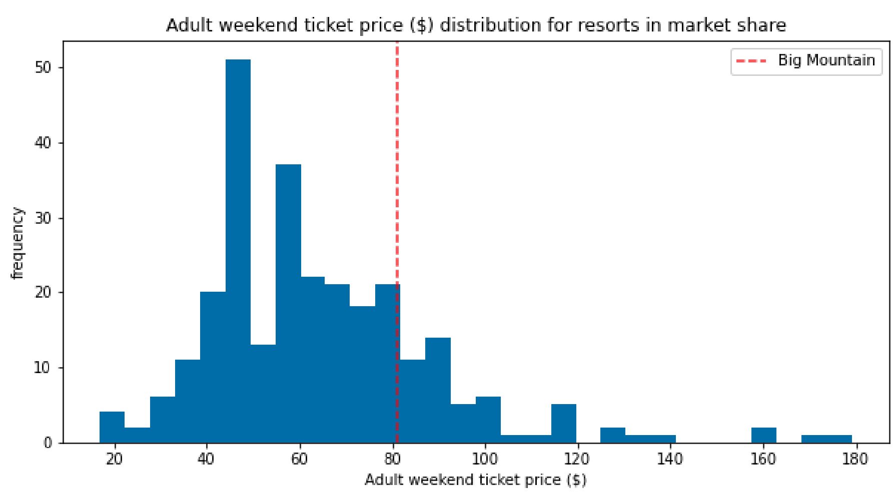
Modeling

After testing several models, a **random forest regressor model was selected for having the lowest cross-validation mean absolute error and the lowest variability.** The following is a brief summary of the steps taken to determine the best model:

1. **A baseline model was developed by taking the average price using the median**. Calculating the mean absolute error demonstrated that using the average would result in being about $19 off.
2. **A simple linear regression model explained over 80% of the variance on the train set and over 70% on the test set**. The linear regression model demonstrated some potential as the result was off the real price by about $9. The result did demonstrate some overfitting, although using the mean did not demonstrate much of a difference from using the median.
3. **A random forest regressor was tested and was found to work the best without scaling and using the median**. The random forest model had a lower cross-validation mean absolute error by almost $1, exhibited less variability, and performance was verified with the test set producing performance consistent with the cross-validation results.

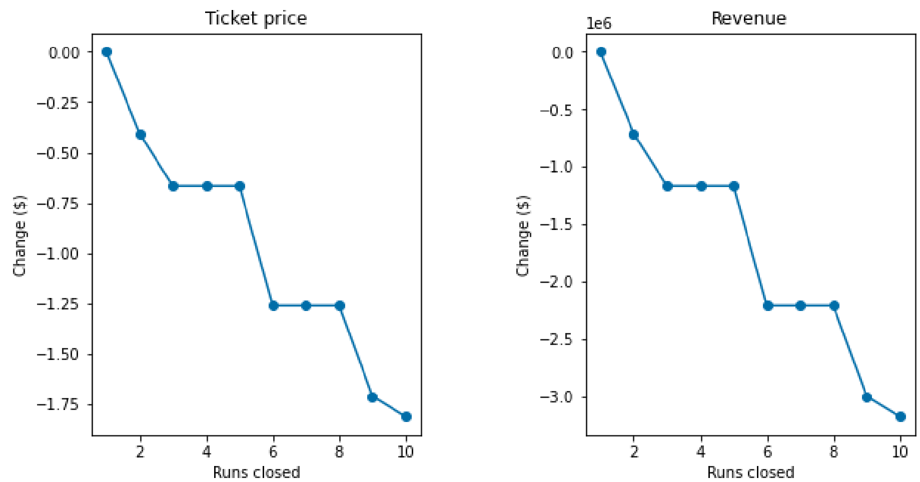
Recommendation

Based on the data analysis in this report, a $95.87 ticket price could be supported in the marketplace based on Big Mountain's facilities (currently, Big Mountain charges $81.00 per ticket).



Given that 350,000 visitors are expected each year, with each visitor skiing five days, the expected increase in revenue would be $26 million. Such an increase would be substantially more than the operating cost for the new ski lift ($1.54 million).

Two potential scenarios are worth further consideration for cutting costs or increasing revenue. First, closing the least used runs could potentially cut costs without decreasing ticket prices substantially - as long as fewer than six runs are closed. It would be worth considering which runs would be closed, and how much revenue would be gained from their closure.



Second, adding a run that would increase the vertical drop by 150 feet and add a chair lift would likely result in a ticket price increase that would be greater than the operating cost of a new ski lift ($3.5 million vs. 1.54 million). Note that if there are additional costs, however, then the difference between ticket price and operating costs would grow smaller.

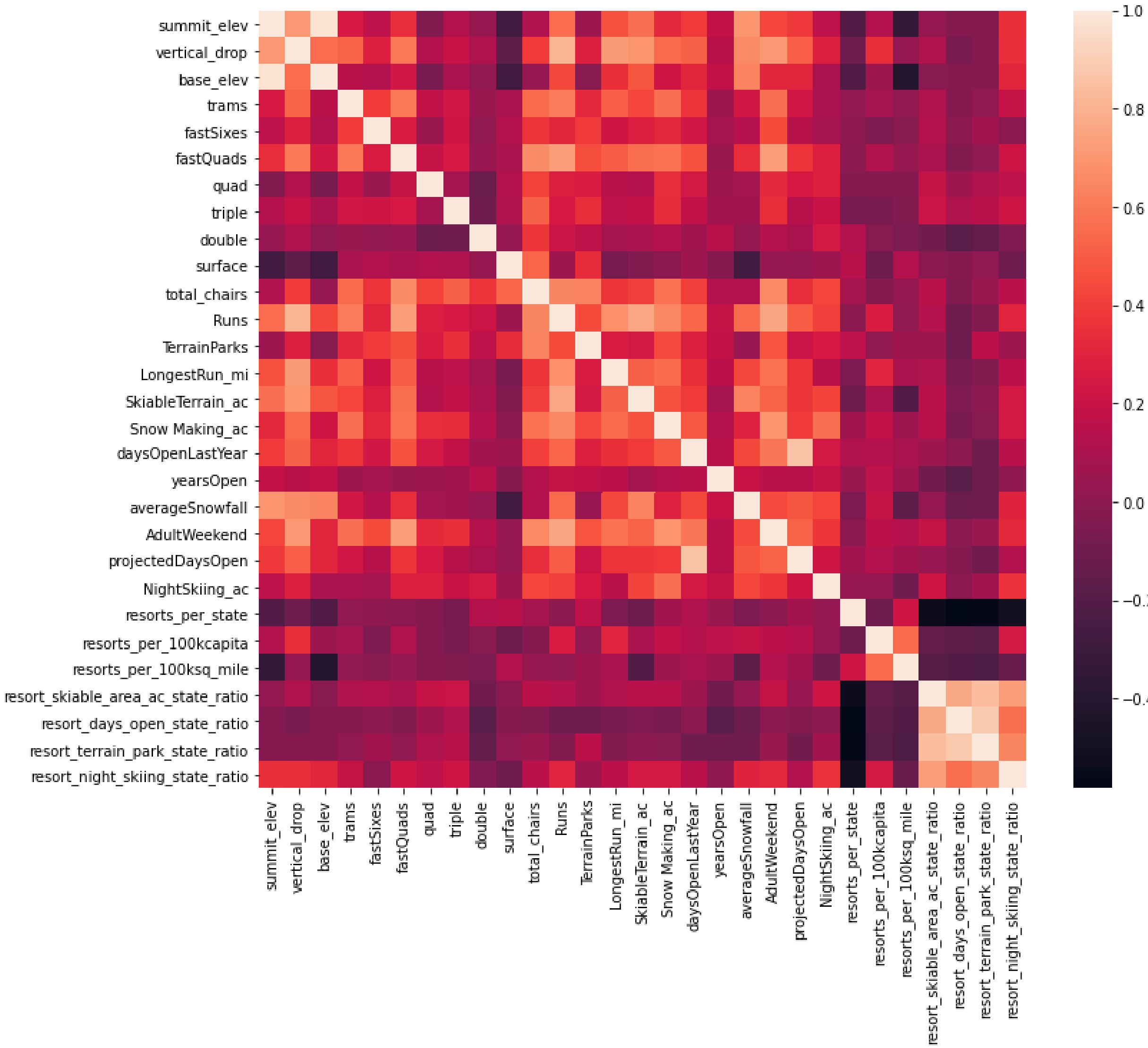
Further Work

For next steps, it would be helpful to consider:

- **Other Costs**. For example, how much is the cost of a new chair? How many years would it take to break even with the purchase of a new chair? What kind of chair is it?
- **Other Revenue**. Are there other ways to increase revenue outside of the ticket prices? For example, does Big Mountain already run a rental business? Given that Big Mountain would be the most expensive resort in Montana, are there non-skiing ways to increase revenue?
- **User Feedback**. Is there any data on which features users value at Big Mountain? Could user feedback be a way to determine if a price increase would be well-received?
- **Additional points of consideration**. What is the expected drop in expected visitors with the new price? How would that impact the increase in revenue?

Appendix A: EDA Visuals

Feature Heatmap



Feature vs. Ticket Price Scatterplots