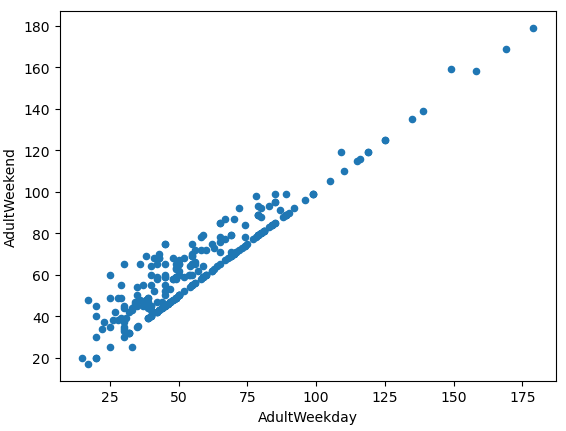
**Big Mountain Data Summary Report**

**Problem Statement** - To determine the optimal ticket price for Big Mountain Resort, by evaluating the impact of different ticket prices on revenue and profitability and evaluating the importance of different facilities on the resort  
  
**Data Wrangling** – Our raw data set (CSV file) had some initial problems. We cleaned the dataset fixing an error in skiable terrain and dropping the "fastEight" column. Rows with missing price data were dropped, retaining available prices. Analysis revealed no instances of weekday prices being higher than weekend prices (Figure 1 below). Population information was merged with the dataset, ensuring complete state coverage. This cleaning resulted in a cleaned data set with normal distributions per feature (Figure 2 below).

 A screenshot of a graph

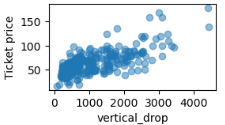
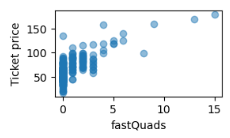
Description automatically generated

Figure 2

Figure 1

**Exploratory Data Analysis –** During this phase of the project it was determined through the use of principle component analysis that there is no clear conclusion that can be made between the ticket price and the state. Therefore, we should consider all states equally moving forward in the analysis when determining the optimal price. We merged the ski data and the state data together and found new interesting relationships which are:

1. Vertical drop, fastQuads, runs, and total chairs show positive correlations with ticket price.

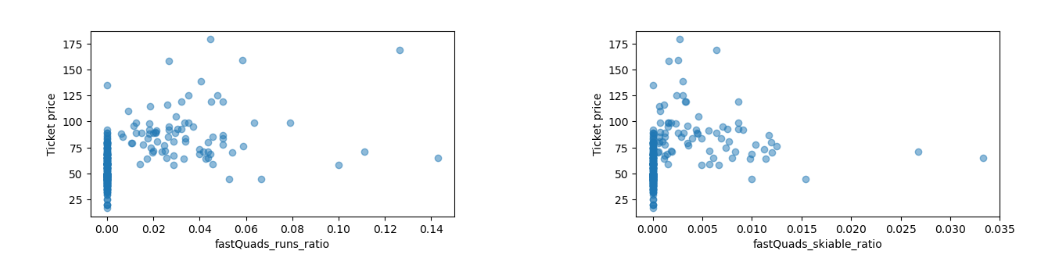
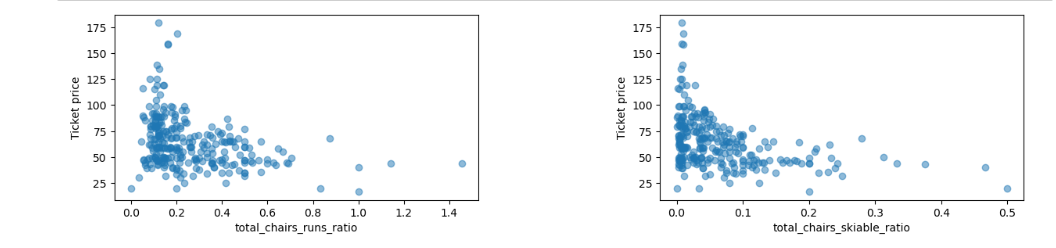


1. Resorts per 100k capita exhibits an interesting pattern where ticket prices may drop before climbing as the number of resorts per capita increases. This suggests that ticket prices could increase with the number of resorts serving a population, indicating a popular skiing area with high demand.

A graph of blue dots

Description automatically generated

1. The ratio of chairs to runs and the ratio of chairs to skiable terrain are introduced as new features, reflecting the ease of transportation within a resort. These ratios can provide insights into how quickly people can access different ski slopes within a resort.



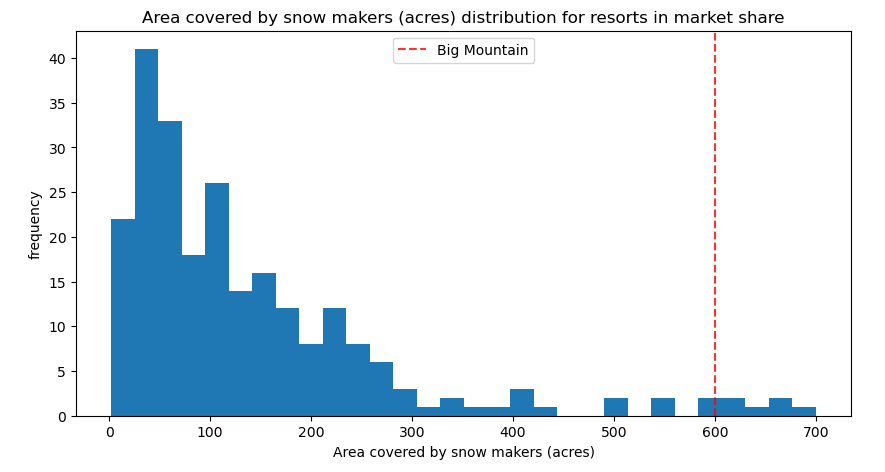
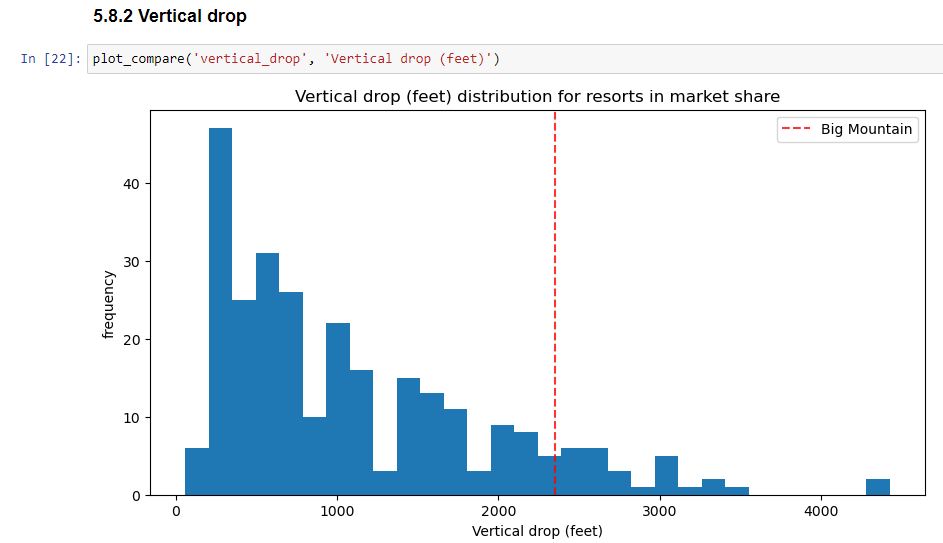
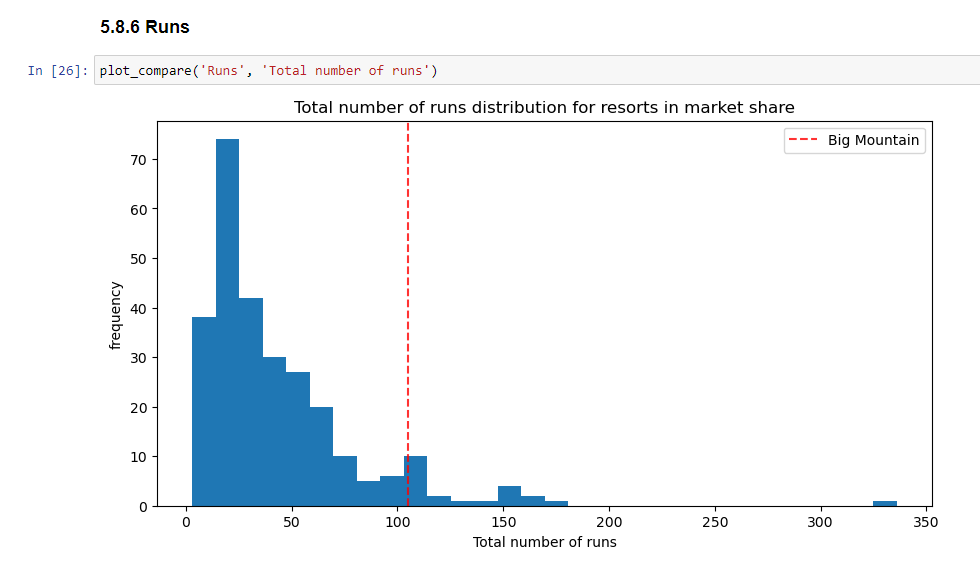
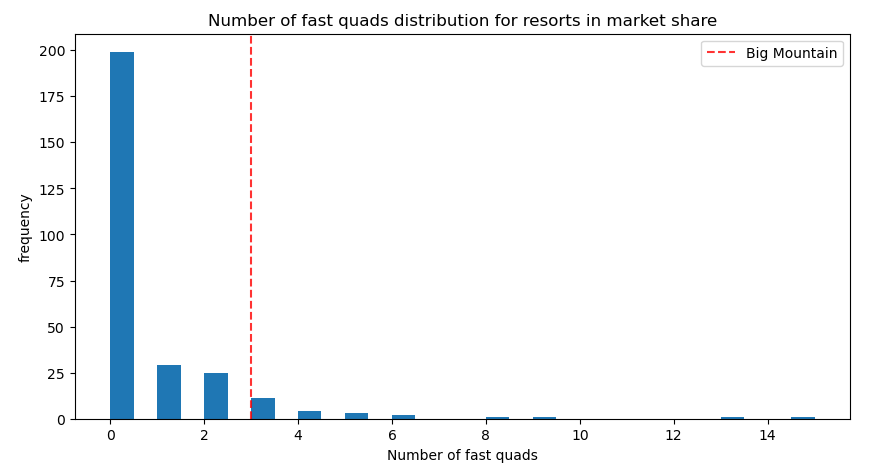
**Model Preprocessing and Algorithms** – Using the mean ticket price, which is $64 as our baseline model. Basic model performance metrices were calculated which determined that just using the mean would give us a model that would be roughly $19 off optimal price. We then created a linear regression model imputing missing values using the mean and median values for their respective columns. The results from this model were similar for both imputing methods and suggested that the model was overfitting. It was determined that not every feature needed to be accounted for in the model and that accounting for the 8 most relevant features would be the most accurate, which are:

* vertical\_drop
* Snow Making\_ac
* total\_chairs
* fastQuads
* Runs
* LongestRun\_mi
* trams
* SkiableTerrain\_ac

We then tested a second type of model called a Random Forest Model. Like the previous model we determined the most relevant features which encouragingly were similar to the above:

* fastQuads
* Runs
* Snow Making\_ac
* vertical\_drop

**Winning Model and Scenario Modeling** - Comparing the two models showed us that the Random Forest Model exhibits less variability and is able to predict ticket prices more accurately than the linear regression model. Therefore, we decided to move forward with the Random Forest Model. When mapping out the most important features, Big Mountain was on the higher end compared to all other resorts (figures below).



For the 4 scenarios presented to us the only one that makes sense to pursue is model 2. It is predicted to increase support for ticket prices by $1.99 and potentially generate around $3,474,638 in additional revenue over the season based on 350,000 visitors per season.

**Pricing Recommendation/Conclusion** – The Random Forest Model suggests that the optimal price based on the current features of the resort compared to other resorts is roughly $96 compared to the current price of $81 with a mean absolute error of $10.39. This means that there is room for an increase in price; however, the validity of the model lies in the assumption that other resorts accurately set their prices according to what the market (the ticket-buying public) supports.

**Future Scope of Work** - There are several deficiencies in the data, including the lack of information on operating costs and visitor behavior, which make it difficult to accurately assess profitability and visitor behavior. The discrepancy between the modeled price and the current price could be attributed to factors such as demand and customer behavior, cost structure, and strategic considerations. It would be important to have a discussion or presentation with the business executives to share the modeling results and gather their feedback. If the model proves useful, it could be integrated into the business's decision-making process for pricing strategy, scenario analysis, and long-term planning. To enable independent use by business analysts, the model could be developed into a user-friendly tool with a graphical interface and proper documentation and training provided to ensure effective utilization.