**Big Mountain Resort Project Report**

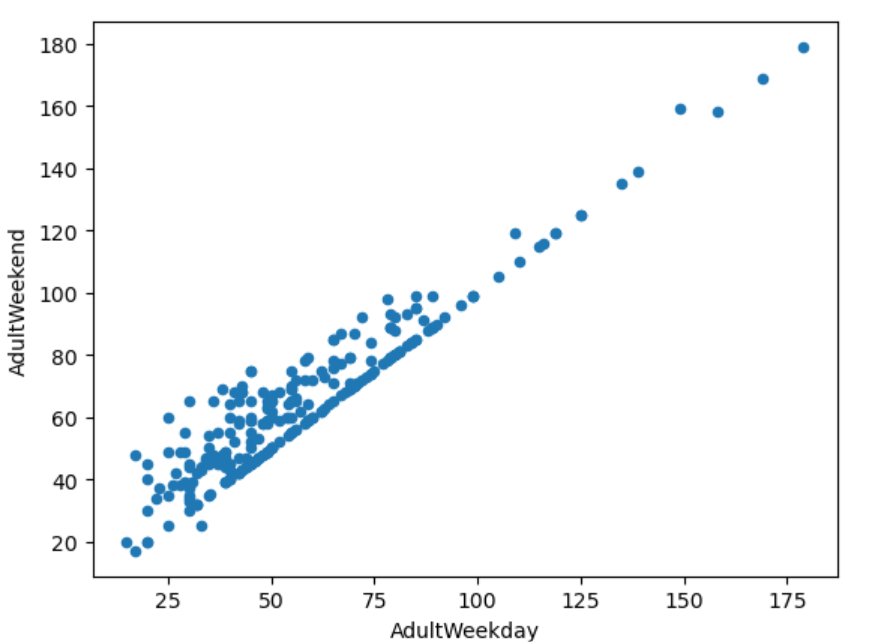
**Problem Statement:**

We stated this Guided Capstone with creating a problem statement for Big Mountain Resort. Big Mountain is a popular ski resort in Montana that offers an expansive number of trails that can be enjoyed by people of all skill levels. They serve a volume of 350,000 visitors every year and they are seeking to serve even more. Their addition of a new chairlift will help them distribute more visitors across the mountain; however, it will increase their operating cost by $1,540,000 this season. While they are seeking to cut costs, their main concern is their approach to pricing. Rather than charging a rate that’s above the market average, they seek a rate that will accurately represent the true value of the experience and amenities they offer.

Our main question/problem at hand is, how can Big Mountain Resort cut costs and adjust their pricing in order to properly capitalize on their facilities; ultimately, to generate as much revenue from it’s 350,000 visitors during the upcoming year?

**Data Wrangling:**

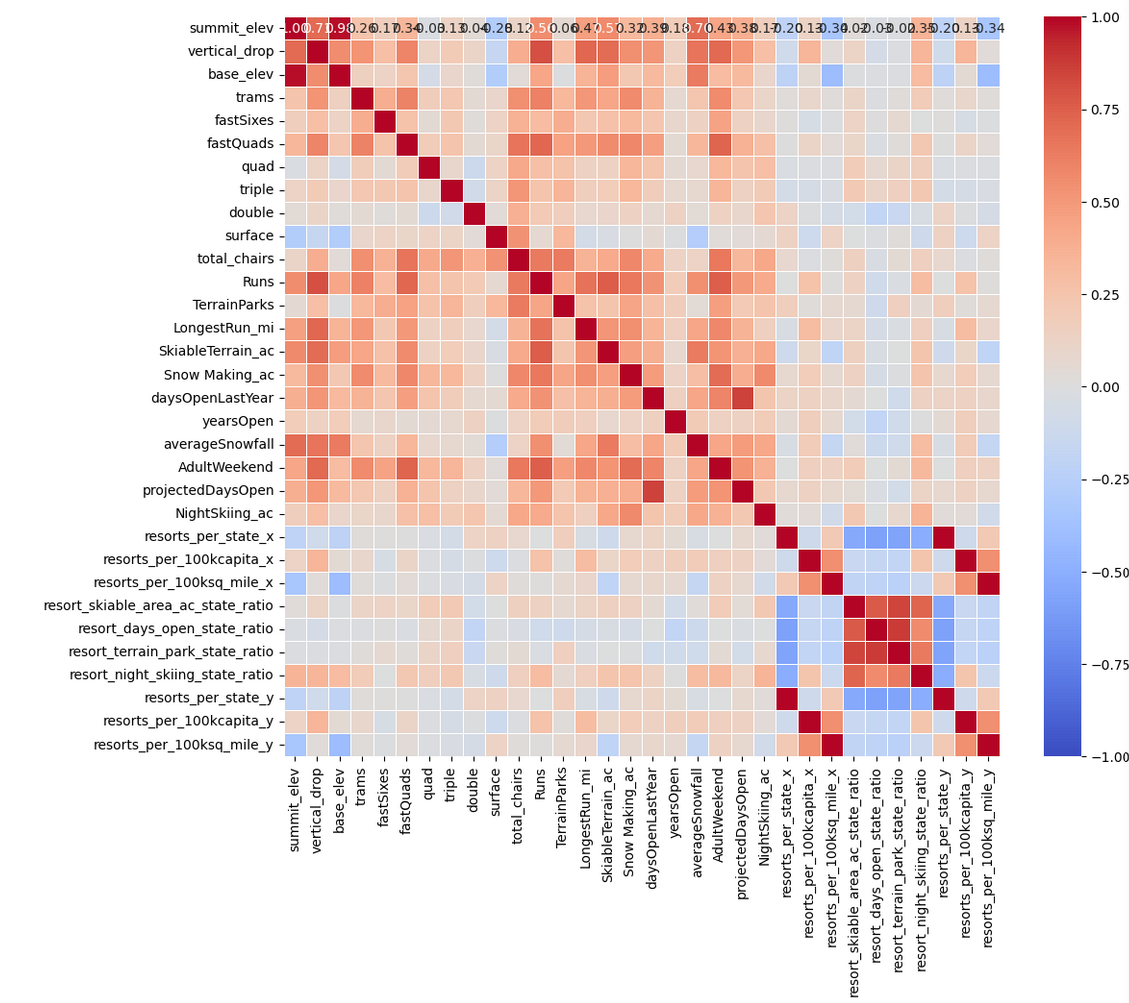
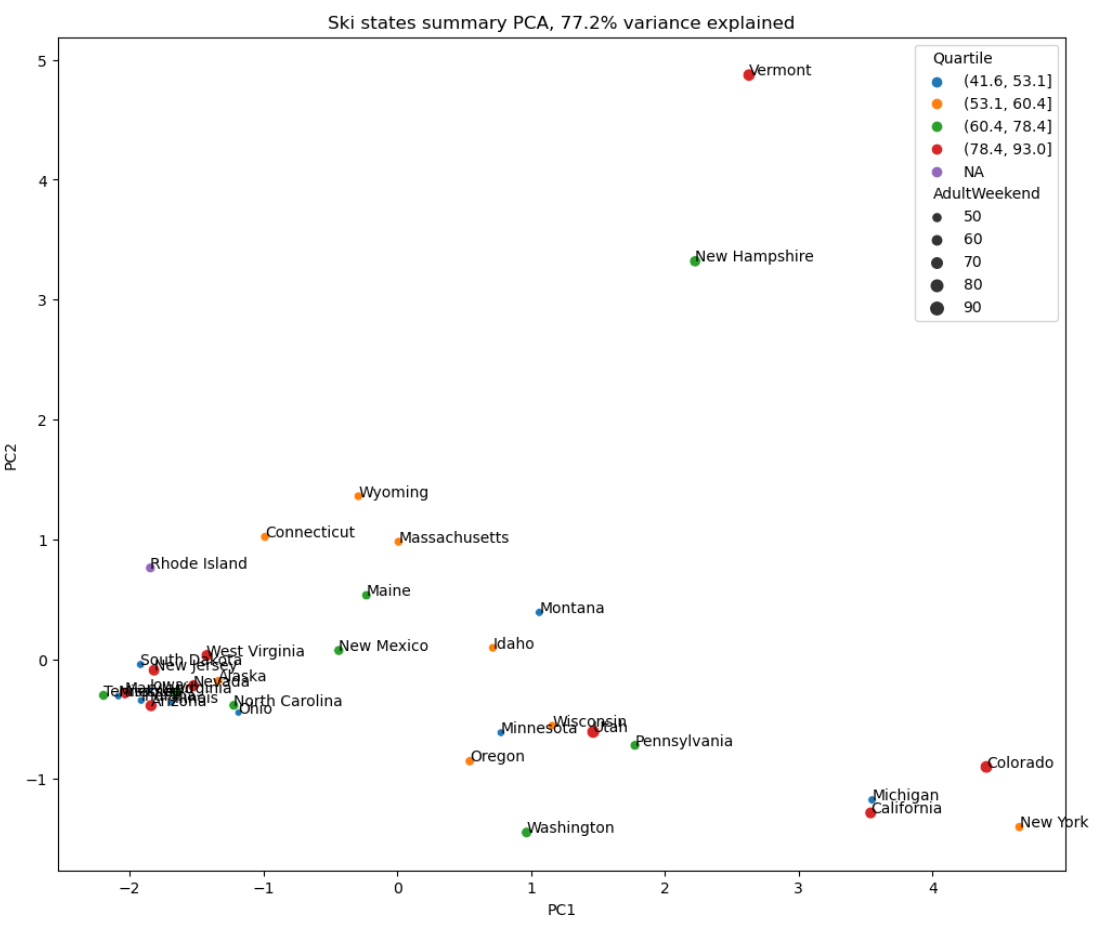
Our Data Wrangling efforts allowed us to clean the data and ultimately determine what our target feature is. After examining both Categorial and Numeric Features. we determined that AdultWeekend is our target feature. The distribution of resorts by Region and State and the distribution of Ticket Price by State helped highlight what the primary target feature should be.



The scatterplot above suggests that weekend and weekday prices are equal, where weekend prices were only higher for resorts that charge less than $100. AdultWeekend prices had the lowest number of missing values when compared to AdultWeekday.

**Exploratory Data Analysis:**

In our Exploratory Data Analysis, we explored a relationship between state level PCA components and the average adult weekend ticket price. We discovered groups of states with similar price levels and features. Our PCA plot included color coded states by quartiles of ticket prices.



Since AdultWeekend is our target feature., we considered the potential of, 'SkiableTerrain\_ac', 'daysOpenLastYear', 'TerrainParks', 'NightSkiing\_ac', 'tota\_charis', 'Runs', and 'fastQuads', to influence 'AdultWeekend'. We used a heatmap to identify highly correlated features, scaling and standardizing features, and creating 'Quartile' to categorize the states.

We created state level features by using state labels. They weren't included as categorical variables. These state level features were aggregated and standardized. PCA displayed the variance across states.

**Preprocessing with Feature Engineering:**

In order to preprocess and train our Data, we started by establishing the baseline performance by using the average price as the prediction for 'AdultWeekend' price. The use of a 'DummyRegressor' with the strategy set to the mean allowed us to achieve this. This baseline model was examined through the use of the training/test sets, and generated train/test versions of performance metrics(R^2Score, MAE, MSE

We then used a pipeline to build a linear regression model. This pipeline included imputing median values,scaling, and feature selection via 'SelectkBest' with 'f\_regression'.

In the initial regression we used SelectkBest with k = 15

1. R^2Score(both train and test) improved greatly compared to the baseline.

2. MAE and MSE (both train and test for each) came back lower compared to the baseline.

3. Mean CV Score had good results with reasonable stability. CV std Dev had more variability

Using 'GridSearchCV' allowed us to determine the optimal number of features(k). Best k was identified and used for the final model and the most impactful features were taken from the coefficients of the linear model.

Next, we tested a Random Forest Regressor with a similar preprocessing timeline that included imputation and scaling. The model used 'GridSearchCV' as well.

The initial Random Forest Model:

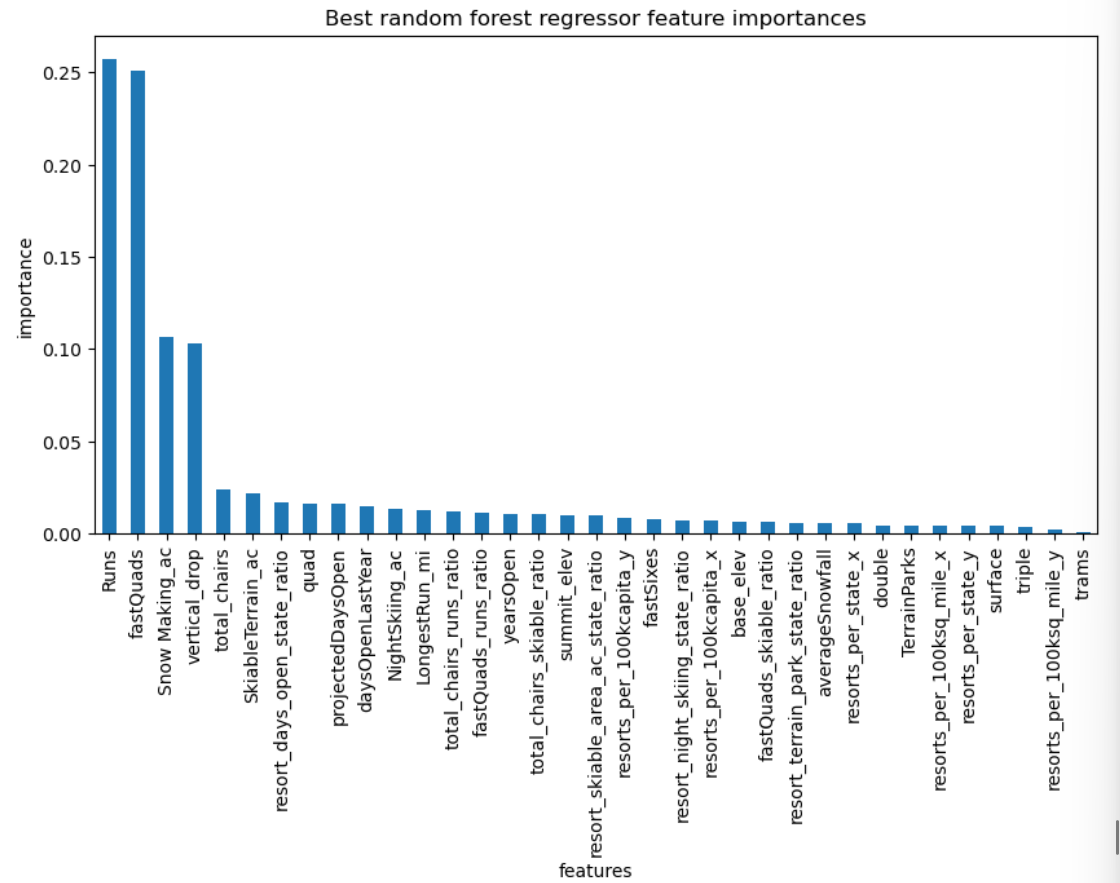
1. R^2Score (both train and test) improved and was competitive with it's variant from the linear model.

2. MAE (both train and test) performance was slightly better than it's variant from the linear model.

3. MSE (both train and test) was lower than the linear model which means better performance

4. Mean CV Score was comparable to the linear model with a bit more stability, and CV Std DEv had lower variability which implies more stable performance

Our grid search gave us the best parameters which enhanced the model's performance, and using a barplot gave us insight into the most important features in the model.



The Random Forest Regressor was chosen as the preferred model because it performed better overall (with a lower MAE and MSE), contained more stable

**Scenario Modelling and Pricing Recommendation:**

Big Mountain Resort is currently charging $81.00 for an AdultWeekend ticket (this is the actual price).

The Big Mountain Resort Modeled Price is $97.96 and the Expected Mean Absolute Error is $10.36 which suggests that there is room for an increase in price.

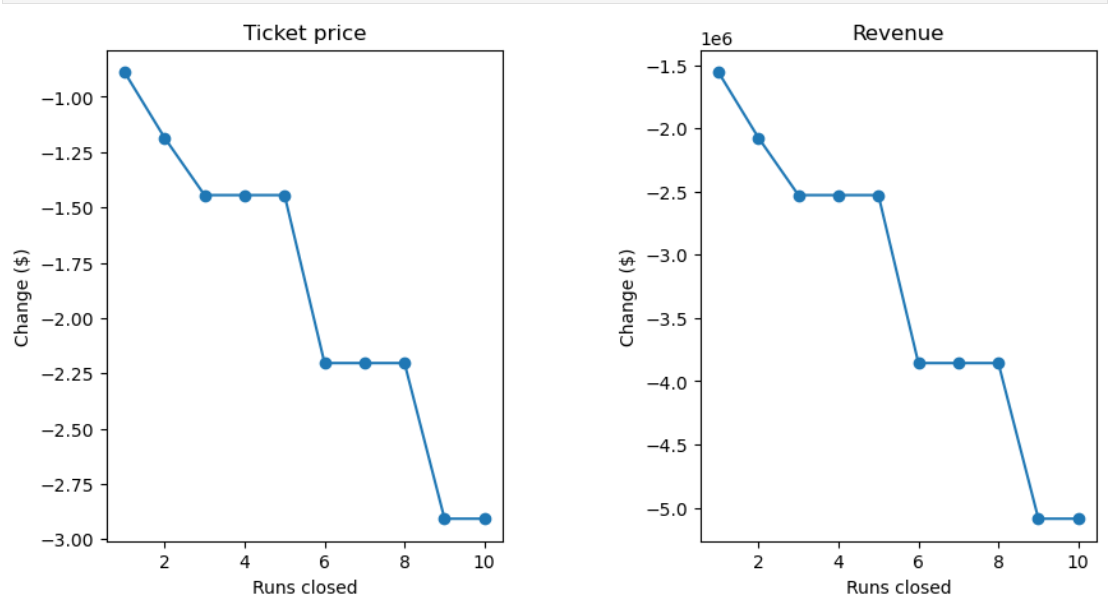
Therefore, the model suggests that Big Mountain Resort could possibly raise their ticket price by about by $16.96 based on the predicted value ($97.96) and the actual value ($81.00).

Our approach to suggesting an increase in price to business leadership could involve: 1. pointing out the mean absolute error to emphasize the reliability of the prediction, 2. Using the histograms created above to show how their price compares to their competition, 3. Highlight the potential revenue increase as a result of the price adjustment and taking in to account the expected number of visitors

Regarding the new chairlift, Big Mountain Resort would incur additional operating costs. We expected there to be 350,000 visitors, and assuming they 5-day tickets, the additional cost of the new chair lift would need to be divided amongst every ticket that is sold. The cost should be factored into the price adjustment if the operating cost is substantial.

We can then use the modeling scenarios below to explore possible growth, cutting costs, and price adjusting

Scenario 1 shows that unless Big Mountain Resort agrees to close ONLY 1 run, they will see a drop in their ticket price, so this is not ideal



Scenario 2 shows that a ticket price increase of $2.22 is sustainable, and would lead to an increase in revenue of $3,888,889 over the entire season if expected visitors remains at 350,000

Scenario 3 shows that adding 2 acres of snow making makes no impact on the potential ticket price increase, nor the revenue over the course of the season. Therefore, adding more snow wouldn't increase revenue and would require more resources and could increase operating costs which would hurt the Resort's bottom line.

Scenario 4 also shows no difference in adding 0.2 miles and 4 acres of snow to the longest run, so it's not ideal.

Scenario 2 is the ideal route to to explore since it allows to an increase in ticket price and revenue

**Conclusion and Future Scope of Work:**

Despite making it this far, we can see that there are several data deficiencies that need to be addressed such as:

1. Operating Costs such as maintenance costs, staff salaries, marketing expenses, utility costs, snow making costs, and all CapEx data.

2. Revenue streams such as any additional revenue earned from equipment rentals, lessons, food/beverage sales, and rooms.

3. Visitor information, such as spending habits, Google and Yelp reviews, Loyalty program members (assuming Big Mountain offers a Rewards program of some sort), and visitor volume based on the month/season/holiday.

Regarding the Price Mismatch issues at hand, it could be the result of the following reasons:

1. Big Mountain Resort could consciously be relying on competitive pricing in order to draw in a wider range of customers with different levels of income, hence the lower price.

2. The resort's perception of quality could be less than favorable. If they have a weak reputation, customers won’t be willing to pay a premium for the resort which would show in a decline in the number of visitors. In response, Big Mountain Resort could have lowered their price to avoid raking a drop in visitor count.

3. A recession or weaker economy could lead to a decrease in customer's spending power. Big Mountain could have lowered their prices to attract and sustain the same number of visitors.

We can gauge the Executive team's reaction to our model/analysis by conducting one on one interviews with them, or administering pulse surveys to get an idea of their understanding/expectations of their current and future pricing strategy.

Business analysts can be empowered to use and explore the model without reaching out to a data scientist by:

1. Providing a user-friendly dashboard that allows a business analyst to input their own parameters and see predicted outcomes

2. Provide training and directives that business analysts can reference when using the model in the future.

3. Ensure that the models updated regularly so that it's accurate.