# Big Mountain’s Project Report:

#### Introduction:

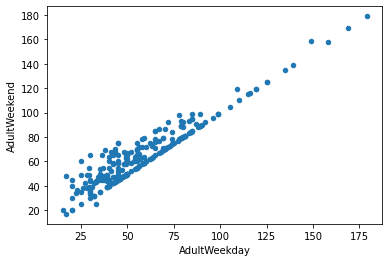
This project was undertaken after a suspicion by Big Mountain that it wasn’t fully capitalizing on its features, and a desire by the resort to gain business insight into what facilities matter most to the customers, which would inform future investment decisions.

#### Problem Statement:

After considering our possibilities, we defined our problem statement as: ‘how can Big Mountain leverage the available data to increase revenue and gain business insight, by the next skiing season?’. Then we continued with our data science process to answer this question.

#### Data Wrangling:

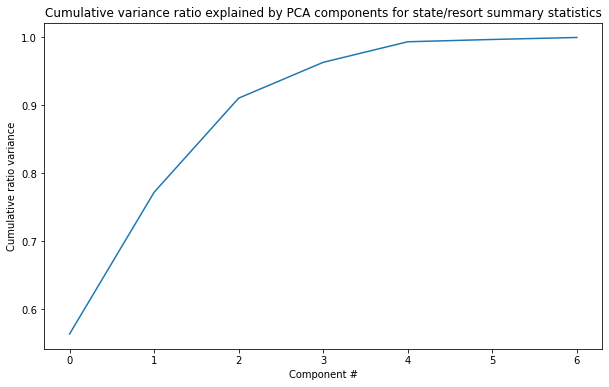
Here our main objective was to figure out if we have the data we need to answer our question, and to identify our target value from the data. We examined the data for missingness, and extracted insights from the data before dropping any missing values. For example we displayed the relationship between the two price categories using a scatter plot:



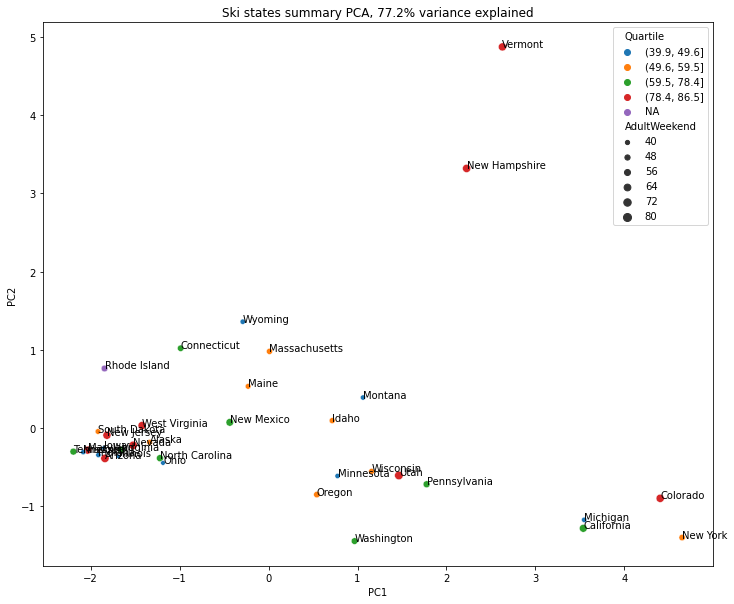
After extracting the possible relevant insights from our data, handling missingness, and treating a few suspect outlier values, we decided to use weekend ticket prices as our target value, since it has less missing data than the weekday ticket prices. We also augmented our data with state populations.

#### Exploratory Data Analysis:

Here, we scaled our data, performed a Principal Component Analysis and plotted the cumulative variance ratio explained by the PCA:



The first four components were responsible for 95% of the variance. We also plotted a heat map that displayed the degree of correlation between features. And we couldn’t find any relationship between price and states (as is evident by the next plot) so we decided to treat states equally:

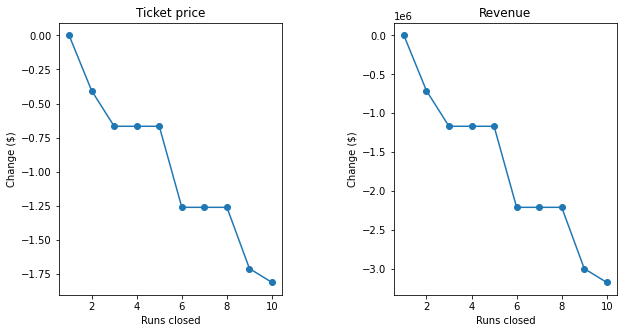


#### Preprocessing and Training The Data:

We developed two models to predict prices here, and we ended up choosing the random forrest model, since it displayed a higher degree of accuracy. We assessed performance using cross-validation. Our model produced predictions that had a MAE of 9.5, which is much better than predicting using the mean which had a MAE of 19.1.

#### Modeling:

We found that our modeled price is $95.87 while our actual price is $81. This shows a lot of room for increasing the revenue just by increasing ticket prices, given the features offered by Big Mountain relative to the market. Another consideration for Big Mountain is closing down runs to cut costs: we recommend looking into the costs that would be cut and comparing them against the lost revenue caused by the closures, a case worth looking into is the closure of 5 runs, since the model doesn’t predict a negative effect between the closure of 3 runs and 5 as is shown in the plot:



Also our model showed that increasing the vertical drop of the longest run, and installing an accompanying chairlift produced a positive increase in ticket price equal to $1.99/ticket, assuming 350,000 visitors, and an average of 5 tickets purchased per visitor, this would amount to an increase in revenue equal to $3,474,638 in the season. Also a small increase in the snow making area in this case had no effect on the ticket price.

#### 