Our firm was tasked to manage and manipulate the provided data from Big Mountain Resort, in order for that company to figure out which facility investments they would like to focus on.  This was done by collecting and cleaning the data. Once this was done, then a model projection could be created. In this instance, the model was tested by a standard train/test split of 70/30(ensures that the model can accurately predict new data).   Before the company can make smart investments in their facilities, one must find the important values that affect both the Weekday and Weekend ticket prices. These vital features that suggests a higher ticket price are vertical\_drop, Snow Making\_ac, total\_chairs, fastQuads,Runs, LongestRun\_mi, trams, and SkiableTerrian\_ac, found after a heatmap was created. However before  the findings sections of this journal,  the data must be reinforced by proof and methods on how the data was obtained.

Step one, cleaning the data.  As with any type of data, the data must be cleaned. After cleaning the data, that was provided, our firm already discover some inconsistencies. These inconsistencies are the elimination of rows containing no data. These rows were dropped with the drop function.  Once this was complete, the amount of rows drop down to 227 from 330 rows. Within the data there was also a discovery that Silverton Mountain  had a skiable area of 26,819ft while on the website there is a claim of 1,819 ft. Because 1,819 ft sounds more reasonable our firm proceeded with the second area of skiable terrain. Now that data is clean, our firm can continue with the data science method which the next step will be exploring the data. This was done by analyzing the ticket prices by using new ratios based on the amount of resorts  per 1,00,000m^2. This will allow the data to ignore the difference of population density between each state. Therefore the data now reflects the amount of resorts supporting a certain area. Then a principal component analysis(PCA) was created to filter through the dimensional data, allowing us to better visualize the data and increase the accuracy of the model. The PCA also would allow the data to be analyze without correlating data. Here on the PCA graph it is seen that the first two components add up to over 75% of variance among the other states.

A picture containing text, line, diagram, plot

Description automatically generated

Now that, the first two components are selected to see if there is a correlation between these components and ticket pricing have a correlation. Even after splitting the ticket prices into separate quartiles, there was no clear correlation on the scatter plot. However, New Hampshire and Vermount do show high ticket prices despite being a smaller area, but with a high resort density. A correlation between the “resorts\_per\_\_100kcapita” and “resorts\_per\_100ksq\_mile,” after a hue\_order argument is added.  Now, there appears to be a type of correlation among the “resorts\_per\_\_100kcapita” and “resorts\_per\_100ksq\_mile,” on the scatterplot. After this, a “components\_” argument is added to allow us to analyze how much each feature contributes to the Cumlative ratio variance. Using this data, our firm can merge these Dataframes the state data with the ski data. A heatmap can be generated to see any correlations of ticket price and features.

A picture containing text, screenshot, pattern, colorfulness

Description automatically generated

The most important features have been identified at this point which are the features that were listed previously at the beginning of this journal. Now, the model can be created and predict the outcome for the four scenarios as well as identifying the returns for the target state which would be Montana.

The four different scenarios are, closing 10 runs, adding an additional run without any new snow coverage, like the second scenario with adding two acres of snow, and increasing the longest run by 0.2 miles to allow advertising of a 3.5 mile run plus the addition of four acres of snow.  Once the model was established it can predict the modelled ticket price which is $100.24 while the Montana resort is charging only $81.00.

The first scenario is vital to Big Mountain Resorts’ success because focusing resources on low used runs would be considered a waste of resources. In the graph below:

A picture containing text, diagram, line, plot

Description automatically generated

It is seen here that the model predicts that there will be initial price drop with the closure of one to two runs. However, what is interesting, is that between the closure of two to eight runs the ticket price reminds steady. The best way to proceed with this data will be analyzing which runs are the least used. Then closing the least popular runs one by one, weeks apart. This way, the resort is given time to analyze true ticket price/revenue drop.

The second scenario is just as worthy, if not more worthy, to investigate further. The model predicts that this would support an increase ticket price of $8.48.  Naturally, this increase of price would raise the revenue ceiling to $14,848,485.  Scenario 3 is identical to the second scenario, except with a difference of increasing snow coverage by 0.2 miles. This minor increase of snow does increase the ticket price by $9.36, bringing a revenue of $16,386,364. The best course of action is establishing this new run with the increase snow coverage.

Lastly, increasing the length of the longest run. Although, skiable terrain, the model predicts that there will be no increase in revenue. Therefore, the other scenarios should be prioritized.

At the end of this study, it is safe to conclude that ticket prices at the Montana resort can and should charge more these graphs will demonstrate that this resort high values of desired skiing factors. Yet compared to other resorts, Big Mountain is only in the middle tier of the ticket prices, suggesting that prices can be raised by implementing scenarios 1-3.

A picture containing screenshot, diagram, plot, design

Description automatically generatedA picture containing screenshot, diagram, plot, graphics software

Description automatically generatedA screen shot of a graph

Description automatically generated with low confidenceA screen shot of a graph

Description automatically generated with medium confidenceA screen shot of a graph

Description automatically generated with low confidenceA picture containing screenshot, plot, diagram, design

Description automatically generatedA screen shot of a video game

Description automatically generated with low confidenceA screen shot of a screen

Description automatically generated with low confidence