Sangeeta Jayakar

Guided Capstone Project Report

Introduction:

The purpose of this work is to create a ticket pricing model in the ski resort market so that the client, Big Mountain Resort, can maximize on their returns. The client suspects their current ticket prices do not reflect their position in the market based on the various features their resort offers. This is likely due to the fact that they do not know which of their facilities customers are likely to pay more for. This project seeks to answer two questions with the analysis. First, based on the facilities offered at Big Mountain compared to other resorts in the market, can Big Mountain change their ticket price to reflect their resort offerings and make them appear more appealing? Second, what facilities can Big Mountain consider investing their resources into in the future?

Data:

The dataset made available to us was a single csv file from the Database Manager, containing entries for each resort. There were 330 rows of data representing each resort, including Big Mountain, with 27 columns representing different features of each resort. The data were almost all numerical, with 2 categorical columns for the state and region of the resort. Features that were depicted included numbers of various types of chair lifts, number of trails, area of skiable terrain, amount of snowfall, area of snow-making ability, vertical drop. The target features were identified as weekday and weekend ticket price. The final cleaned dataframe used for further analysis contained 277 rows after dropping resorts that did not have values for “AdultWeekend” ticket price, the target feature, and 25 columns after dropping “AdultWeekday” ticket price and “fastEights'' due to missing values. An additional table of data was acquired from Wikipedia that contained population and area data for each state1. These statewide summary data were used to relate the ski resort data in the context of population density and resort density (Figure 1).

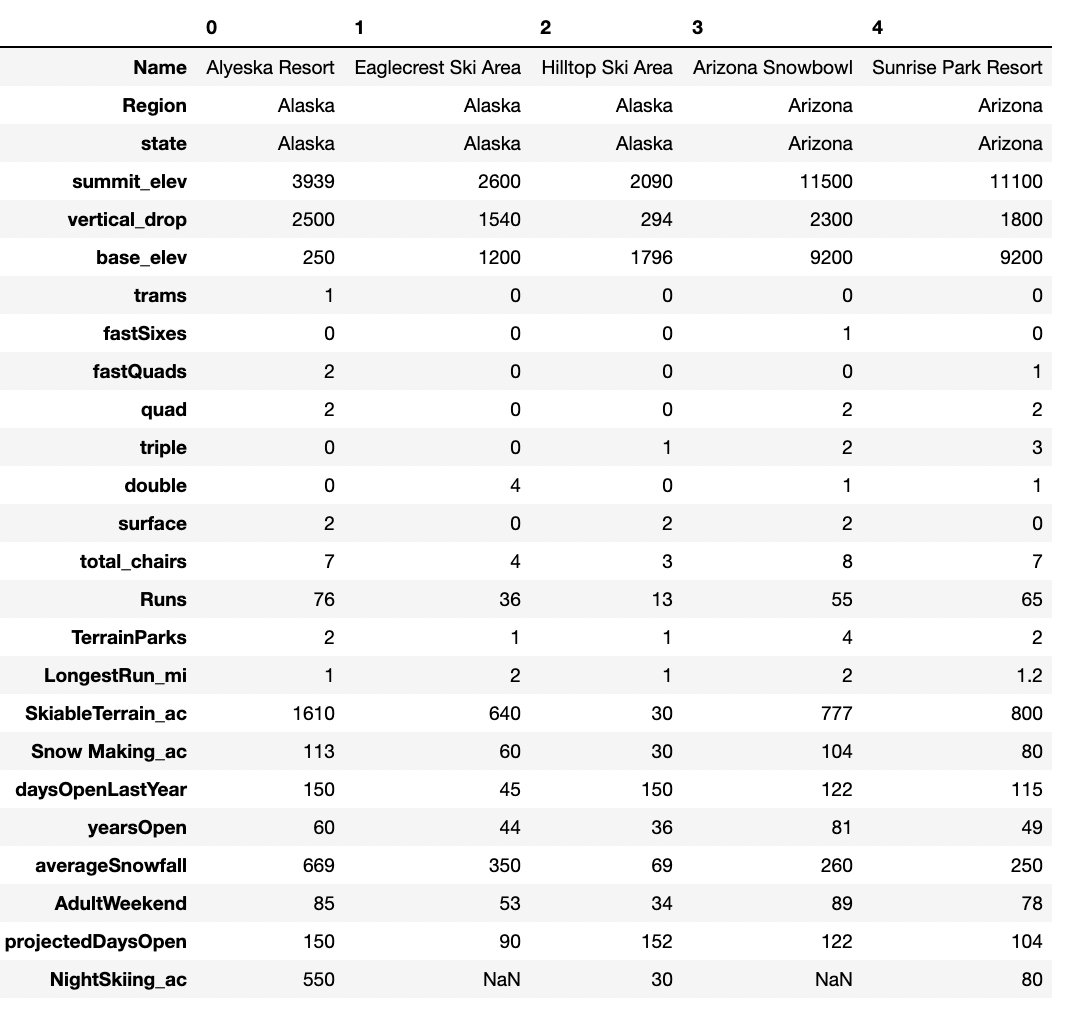
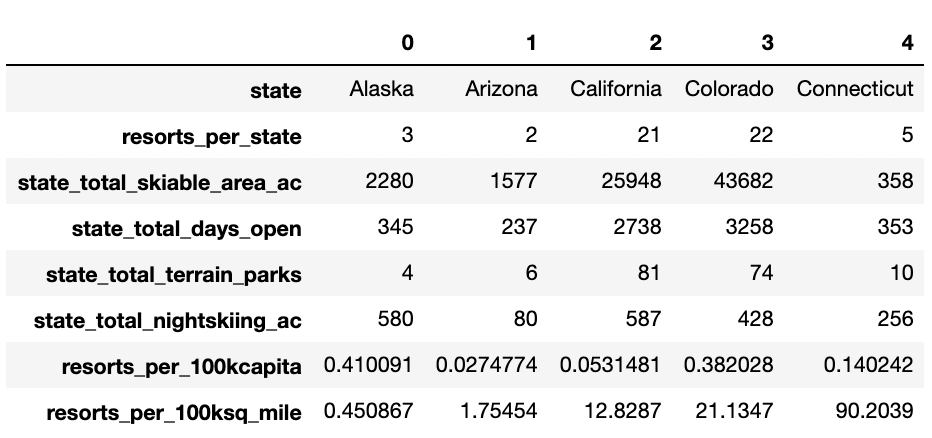
 

Figure 1. Transposed DataFrames showing first 5 ski resort entries (left) and first 5 state summary entries (right).

Methods:

Exploratory data analysis was performed on the ski resort and state summary data sets. The goal of this approach was to assess the features of the ski resorts in relation to the population and area of the states they are in. Data were scaled so that Principal Components Analysis could be performed. Resorts per 100K capita and Resorts per 100sq mile were the two features that seemed to contribute most to the ticket price. Features related to area, such as number of runs and number of chair lifts to transport customers to the runs. Also features that correlated with ticket price such as night skiing, and snow-making acres were also thought to be important to consider when modeling. All states were treated equally for the modelling. After merging the ski resort and state summary datasets, data for individual resorts in relation to the data for its state were calculated so that data for resorts in different states could be compared to each other.

The heat map (Figure 2) and scatter plots (Figure 3) show a positive correlation between the ratio of area for Night Skiing with the number of resorts per capita. When resorts are more densely located with population, more night skiing is provided for these resorts. Also, looking at the target feature of ticket price, labeled in the heat map as “AdultWeekend”, it is apparent that several other features are positively correlated. FastQuads, Runs, and Snow-making ability (this one indicates that customers are more likely to value the concept of guaranteed snow, which is obviously required for skiing), which is more positively correlated than “skiable terrain area”. So total skiable area is not as important as guaranteed snow on the skiable terrain. Another interesting correlation is the features “Runs” and “total chairs” with ticket price. Vertical drop also seems to be a factor that is correlated with ticket prices.

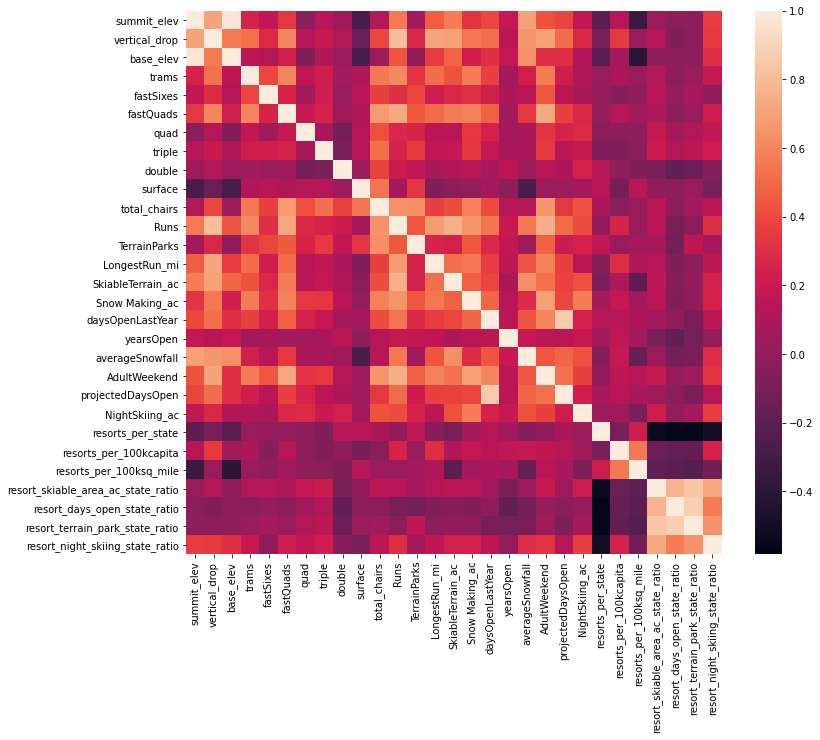
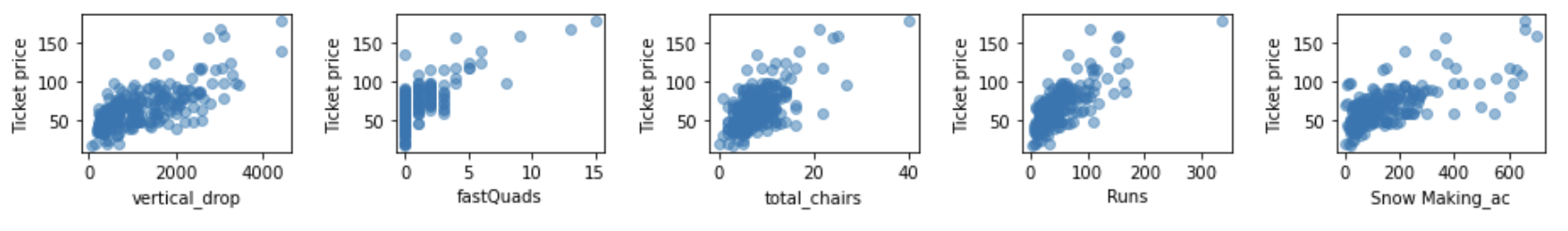


Figure 2. Heat map of correlated features from merged ski resort and state summary datasets.

Figure 3. Scatter plots of features of interest correlated with ticket price.

Analysis:

Both Linear Regression and Random Forest models were tested. A pipeline was built to impute missing values, scale the features, train the model, and calculate the model performance. This pipeline was applied to test a linear regression model, and then a random forest regressor model. For random forest, it was determined that the median would be used to impute missing values, but the features would not be scaled. With this model, the important features were fastQuads, runs, snow-making acres, and vertical drop (Figure 4). This model has a lower cross-validation mean absolute error by almost $1, and has less variability.

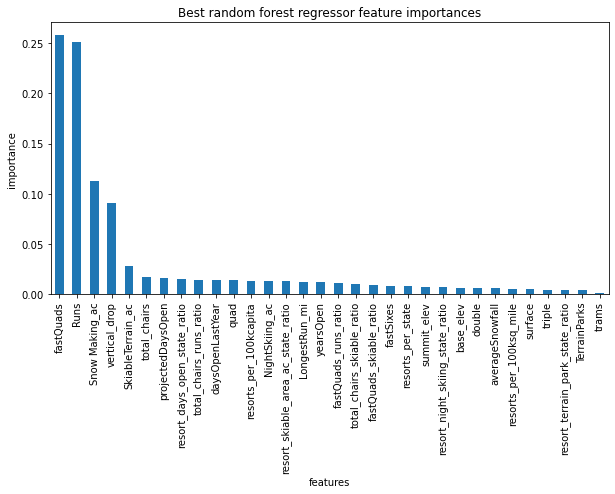


Figure 4. Important features based on Random Forest Regressor model.

After the Random Forest Model was selected, we tested what the price of a ticket for Big Mountain should be based on the current offering of facilities. We then tested several different scenarios based on what facilities could be invested into in the future.

1. Permanently closing down up to 10 of the least used runs.
2. Increase the vertical drop by adding a run to a point 150 feet lower down but requiring the installation of an additional chair lift to bring skiers back up, without additional snow making coverage.
3. Same as number 2, but adding 2 acres of snow making cover.
4. Increase the longest run by 0.2 mile to boast 3.5 miles length, requiring an additional snow making coverage of 4 acres.

Results:

According to our Random Forest Model, Big Mountain Resort should currently charge $95.87 instead of $81 for their ticket price, based on the facilities that they offer compared to other resorts in the ski resort market. This model was also used to test four different scenarios that were proposed by Big Mountain Resort. Scenario two, which aimed to add an extra 150m to the vertical drop by adding one more run along with one additional chair lift to accommodate this new run point, would support the increase of the ticket price by $1.99. Over the season this would amount to revenue of $3,474,638. In addition to this, closing down one of the least used runs as proposed in Scenario One, would not decrease the revenue of Big Mountain, and could thus help to cut down on operation costs.

Conclusions:

This work aimed to answer the question of how Big Mountain can change their ticket price to reflect their facilities offered compared to other ski resorts in the market, as well as what facilities can be invested into for future growth. Big Mountain Resort should change their current ticket price to $95.87, and can further raise prices by an additional $1.99 if they make the recommended changes proposed.

References:

1. List of U.S. States, Wikipedia <https://simple.wikipedia.org/w/index.php?title=List_of_U.S._states&oldid=7168473>