Big Mountain Ticket Pricing Strategy: Data Analysis, Modeling and Recommendations

Created by: Marin Stoytchev **Date:** Oct. 9, 2020

Abstract: Using the dataset provided, we built a Random Forest regression model for predicting the ticket prices for ski resorts in the USA. The model has a mean absolute error of $10.38 which is approximately 16 % of the average ticket price of $63.81 and 13 % of the current ticket price of $81.00 for Big Mountain resort. The modeled ticket price for Big Mountain is $94.22 which suggests the possibility of increase of the current price. Two cost-saving and revenue-increasing scenarios are proposed. Before moving forward with the implementation of any of these scenarios, we propose further data discovery and analysis, and model tuning using data for resorts which are direct competitors of Big Mountain resort.

Keywords: USA ski resorts, ski resort ticket prices, ski resort facilities, Random Forest regression, predictive analytics

Introduction: Determining the optimal ticket price for a ski resort is essential in deriving the most profit for the resort, while remaining competitive in the respective market segment. Techniques of data science and predictive analytics can be used to predict optimal ticket prices and to derive optimal strategy for reducing operational costs and increasing profit margins. In this project, we are provided with the ***ski_resort_data.csv*** dataset which consists of 330 entries (rows) and 27 columns with the rows representing different ski resorts in the USA and the columns representing different features of these resorts, respectively. After cleaning the raw data and augmenting it with population and area numbers for each state the final data used for modeling was saved as ski_data_cleaned.csv in folder data.

Project Objective: The goal of this project is to use techniques of data science to provide optimal ticket price for Big Mountain resort and to evaluate different strategies (scenarios) for reducing cost and increasing profits for the resort, while remaining competitive in the respective market segment.

Exploratory Data Analysis: The dataset was imported in a jupyter notebook and data wrangling, statistical calculations, and visualization were performed using Python. We plot the following figures from the cleaned data:

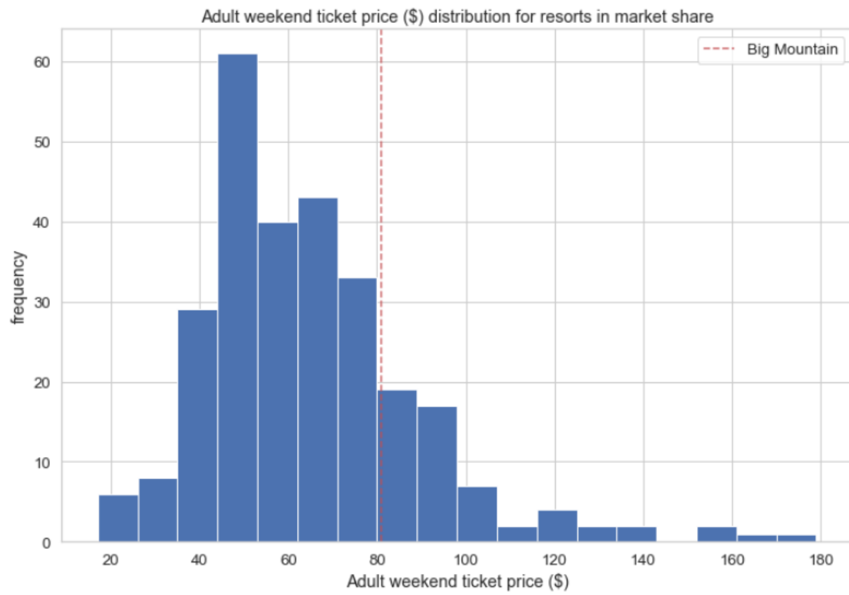

Fig. 1: Distribution of ticket prices for resorts in market segment.

Fig. 1 shows the histogram of ticket prices for all ski resorts in the market segment. The ticket price for Big Mountain is in the top 20 % of the ticket price range.

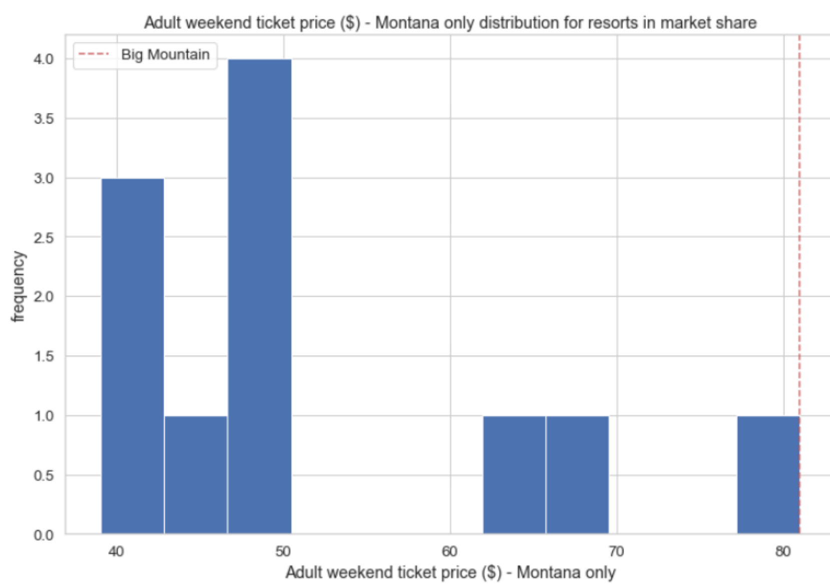

Fig. 2: Distribution of ticket prices for resorts in the state of Montana only.

As can be seen from Fig. 2, Big Mountain has the highest ticket price for ski resorts in Montana.

Figures 3-6 show the Big Mountain position in the market segment regarding the most important resort features as identified by our model.

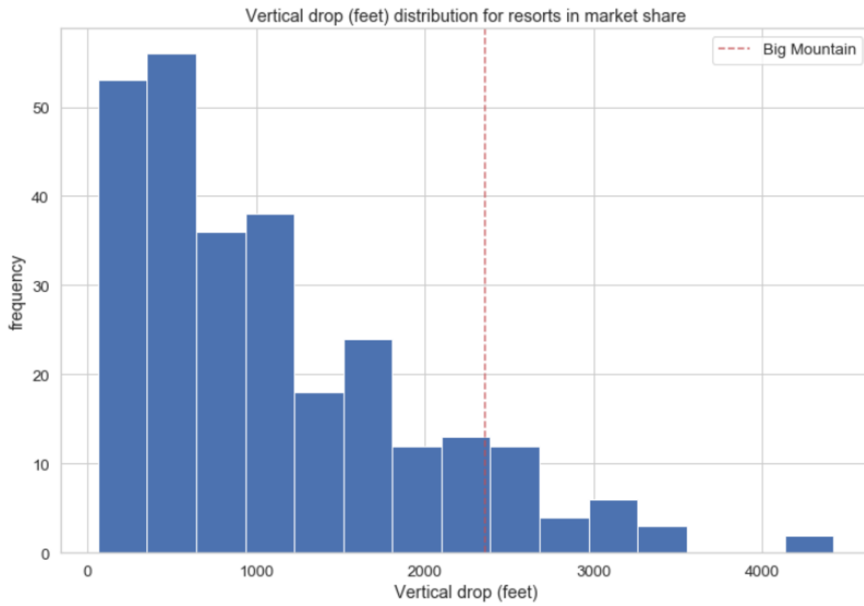

Fig. 3: Distribution of vertical drop for resorts in market segment

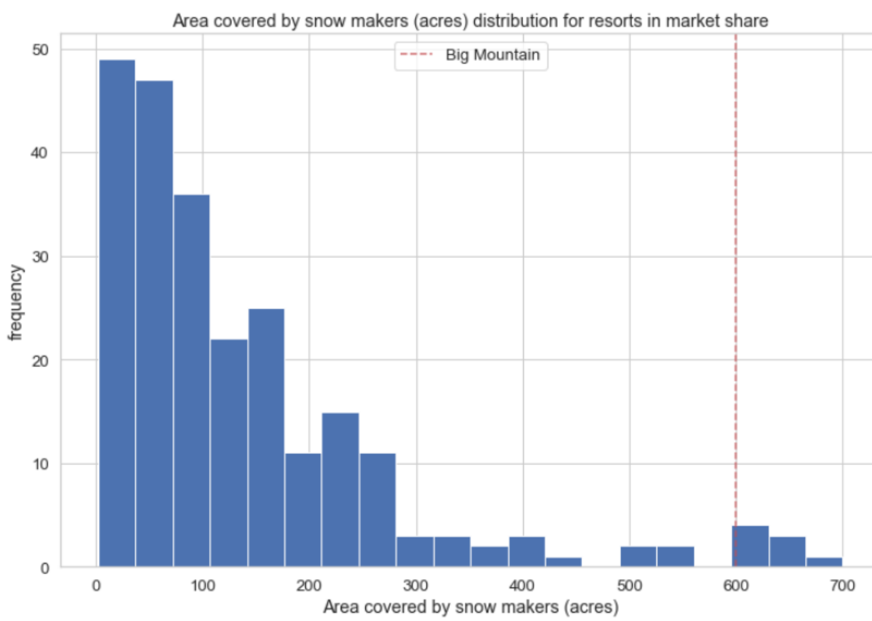

Fig. 4: Distribution of snow making area for resorts in market segment

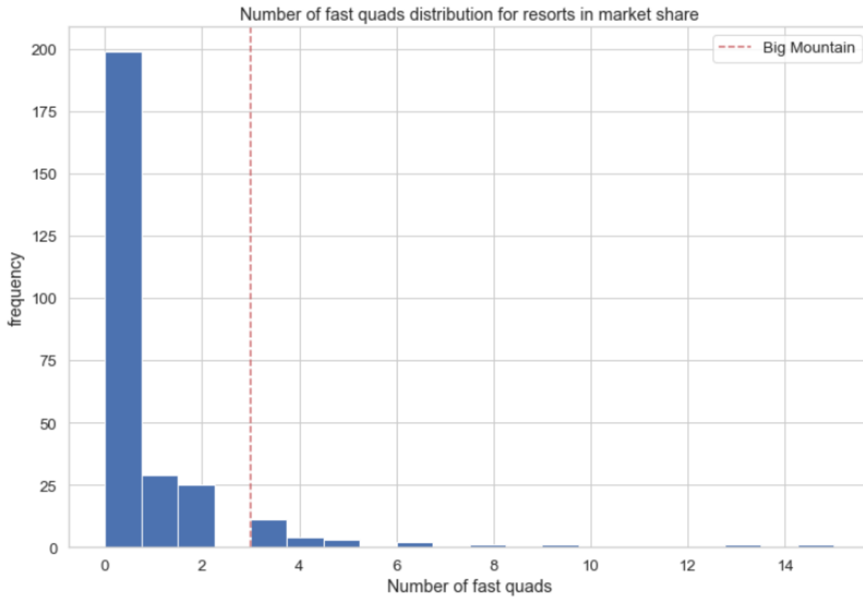

Fig. 5: Distribution of number of fast quads for resorts in market segment

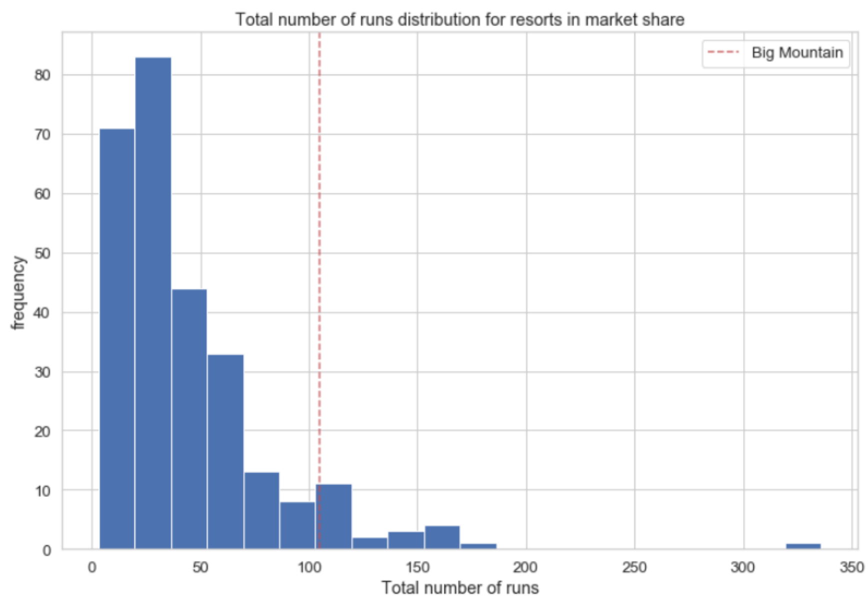

Fig. 6: Distribution of total number of runs for resorts in market segment

All distribution plots of the most important features show that Big Mountain ranks very high in all categories.

Model Selection: In this project, we have built and optimized two predictive models: a) Linear Regression model and b**)** Random Forest Regression model. After optimization using cross-validation the best Random Forest model is found to have the smallest error. In addition, the errors obtained for both the training and the test set are very close which indicates that the model generalizes well and is expected to make reliable predictions on unseen data. The selected model was saved as ***"model.pkl"*** file for future use.

Figure 7 shows the feature importance in determining the ticket prices of resorts in the market segment according to the best performing Random Forest Regression model. According to the model the most important resort features are: fast quads and runs with almost equal weight close to 0.25, and snow making area and vertical drop which are significantly far behind with approximate weights of 0.11 and 0.09, respectively.

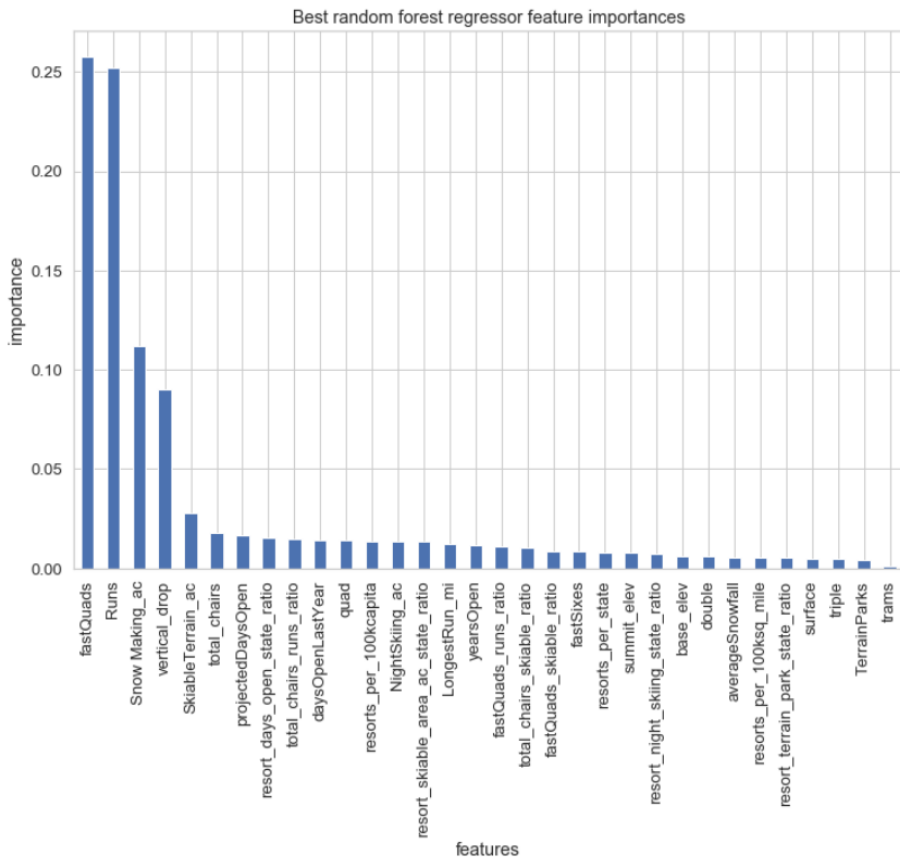

Fig. 7: Feature importance derived from best Random Forest Regression model

Predictions: The selected model has been trained on the entire data set (excluding Big Mountain data) and prediction of the ticket price for Big Mountain has been made. The modeled ticket price is $94.22 which is significantly higher than the current price of $81.00. Although the difference between the modeled and the current price is close to the mean absolute error of the model of $10.38, the results from the model suggest the possibility of price increase by almost 10%. However, at his point we do not recommend such increase before additional data gathering, analysis and model tuning which will be discussed in the Conclusions section.

Several cost-cutting and profit-increasing scenarios have been explored with our model. In order to estimate the revenue loss/gain in each of these scenarios we have assumed that the expected number of visitors over the season is 350,000 and, on average, visitors ski for five days. Out of the explored scenarios, the two most viable and promising are presented below.

Scenario 1: Closing up to 10 of least used runs. In this case, we are looking for operational cost savings as opposed to increase in revenue. As illustrated by the plots in Fig. 8, closing one run (the least utilized one) does not require change in ticket price and, thus, does not result in revenue loss. We cannot provide here a figure for the amount saved since we do not have data for the operational costs for the resort facilities.

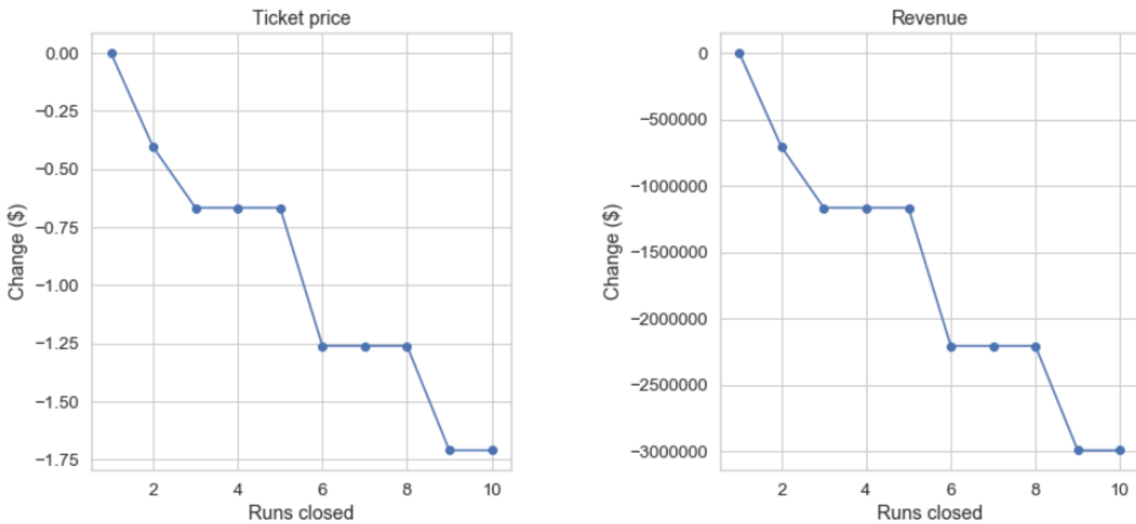

Fig. 8: Changes in ticket price and revenue with number of runs closed.

In case the management decides to pursue more aggressive strategy in closing multiple number of under-utilized runs, the model predicts that up to five of these runs can be closed with relatively small penalty in ticket price and revenue. However, in order to decide if this strategy is profitable trade-off analysis of revenue vs savings in operational costs must be performed.

Scenario 2: In this scenario, Big Mountain is adding a run, increasing the vertical drop by 150 feet, and installing an additional chair lift. This scenario supports increase of $1.99 of the ticket price which is expected to bring over the season additional revenue of approximately $3.5 mill. As with the previous case, trade-off analysis of revenue vs increased operational costs and capital needed for building the additional facilities must be performed.

Conclusions: Using the available data we have performed analysis and built a model for predicting the optimal ticket price for Big Mountain. The model predictions support increase of the current ticket price by up to $94. In addition, different scenario cost-cutting and revenue-increasing scenarios have been explored. The two scenarios with the most profit potential have been presented above. As the immediate safest way of increasing profit, we recommend that the least utilized run be closed without change in current ticket price.

In order to make further recommendations, we believe that additional data analysis and model tuning are required. If such data is available, we recommend exploring the visitor profile in terms of in-state vs. out-of-state. Depending on the results of this analysis, we believe that a model which focuses on ski resorts in the neighboring states of Washington, Oregon, Idaho, Wyoming, and Utah, and the state of Montana would provide more accurate predictions and would help in developing the most profitable strategy. In addition, besides the ski resorts in the states mentioned above, it is highly likely that ski resorts near Calgary and Vancouver, Canada, might need to be included in the analysis and modeling.

Note: Python code for performing data analysis and modeling can be found at https://github.com/marin-stoytchev/DataScienceGuidedCapstone.